Essays on Family Economics and the Macroeconomy

Lidia Cruces De Sousa

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

Universidad Carlos III de Madrid

Advisor:

Luisa Fuster

May 2023

A mis padres, a mi hermana y a Xabi

Acknowledgements

I am indebted to my academic parents, Luisa Fuster and Andrés Erosa, for their invaluable support and guidance. Luisa, your remarkable patience, countless hours spent meticulously reviewing my work, and our lengthy meetings mean the world to me. Andrés, I am immensely grateful for your infectious optimism and the way you consistently challenged me to elevate the quality of my papers.

A special acknowledgment goes to Jaime Alonso, whose belief in my potential as a researcher has brought me to this point of writing my dissertation acknowledgments today. I am also very thankful to Timothy J. Kehoe, my wonderful host at the University of Minnesota. Thank you for your support and invaluable advice. I especially thank him and Manuel Amador for invaluable guidance and insights to improve my research. I am also grateful to Mariacristina De Nardi for her helpful comments.

I sincerely thank the outstanding faculty members at Universidad Carlos III. In particular, I extend my thanks to the remarkable macro group— Antonia Díaz, Juan José Dolado, Carlo Galli, Belén Jerez, Matthias Kredler, Evi Pappa, Clara Santamaría, Hernán Seoane, Emircan Yurdagul, and Felix Wellschmied— for their comments and help throughout my Ph.D. Moreover, I would also like to thank Angélica, Arancha, and Mercedes for their tireless efforts in providing administrative support, always going above and beyond to assist us.

Throughout this enriching journey, I had the privilege of learning from my exceptionally talented coauthors: F. Javier Rodríguez, Isabel Micó-Millán, and Susana Párraga. Their expertise and collaborative spirit have contributed significantly to this thesis; half is theirs. Javier and Isabel have also been my Ph.D. colleagues. Javier, thank you for being your authentic self and friend. Isabel, I appreciate the meaningful moments we have shared and your support as a friend.

I am grateful to my office mates and dear friends, Florencia Airaudo and Agustina Martínez. Our journey together has been truly remarkable. Thank you for being there during the best times and, more importantly, during the challenging moments. You celebrated every milestone in my career and offered solace and support during difficult times. *Gracias chicas, por ser mis bruxas favoritas.*

I am also grateful to the boys in my cohort, Álvaro Delgado, Morteza Ghomi, Telmo Pérez, Henry Redondo, and Fernando Riveiro. Thank you for your support, encouragement, and friendship throughout this challenging process. Besides, I would like to extend my gratitude to Diego Malo, Iván Acosta, Belén Rodríguez, Rubén Veiga, and Ismael Gálvez for the memorable moments we enjoyed together.

Thank you, Nicolò Ruso, my dear Italian friend, for your incredible generosity in opening the doors of your home in Minnesota and sharing your office space with me. Our connection and friendship were instantaneous, and I cherish the memories we created together. I also want to express my heartfelt gratitude to the "European crew" in Minnesota, Marta Morazzoni, Marta Cota, Katherina Thomas, and Ante Šterc. Thank you for making my stay extraordinary. Additionally, I sincerely thank Jiyeon Hwang for creating a comforting sense of home and our enduring friendship since then.

Thank you, Raquel Justo and Adrián Mérida, for cheering me up, supporting me, listening to me, and giving me advice. Our bond is truly indescribable, and I am grateful for our friendship. Thank you for everything.

Last but not least, outside academia, I thank my family and friends for always supporting me and remembering me the priorities in life. *A mis padres, Miguel y Emma, por educarme en un entorno repleto de amor y respeto. Gracias por sacar siempre lo mejor de mí. A mi hermana, Rocío, por ser mi media naranja, la persona a la que más admiro. Gracias por ser tan valiente y por demostrarme que la vida hay que vivirla. A mi compañero de vida, Xabier, por tu compañía, tu cariño y tu respeto. Gracias por hacerme feliz. A mis padrinos, por ser mi familia en Madrid y acogerme en su casa como una hija. No podría haberlo hecho sin vosotros. Os quiero.*

Published and submitted content

The empirical work in Chapter 2 of this dissertation (coauthored with Francisco Javier Rodríguez Román) is also included in our joint project "Quantifying the Impact of Childcare Subsidies on Social Security", which comprises Chapter 2 of the Ph.D. thesis "Essays in Family Economics" by Francisco Javier Rodríguez Román (2021). Programa de Doctorado en Economía por la Universidad Carlos III de Madrid. URL: http://hdl.handle.net/10016/33040

A summary of Chapter 3 of this dissertation is included in Lebedinski et al. (2021).

Chapter 4 of this dissertation (coauthored with Isabel Micó Millán and Susana Párraga) has been sent for review to be published as *Documento de Trabajo del Banco de España.*

Chapter 4 of this dissertation (coauthored with Isabel Micó Millán and Susana Párraga) is partly included in Chapter 2 of the Ph.D. thesis " Essays on Household and Public Finance" by Isabel Micó Millán (2023). Programa de Doctorado en Economía por la Universidad Carlos III de Madrid.

Summary

There is a growing interest among macroeconomists in modeling the endogenous decisions of women and incorporating explicit heterogeneity using rich microdata to discipline their models. In this way, we can understand the heterogeneous impact of government policies that might be potentially different from those obtained using a representative agent model. Motivated by the persistent gender disparities in the labor market, household roles, and portfolio decision-making, my dissertation is structured around four chapters that delve into these issues while explicitly modeling women's decisions.

In the first chapter, *A Quantitative Theory of the New Life Cycle of Women's Employment*, I develop a theory to provide a unified explanation for the employment and fertility changes across cohorts of college-educated married American women. The employment profile of cohorts born before the fifties was initially flat and rose steeply after age 30. In contrast, women born after the mid-1950s started with a much higher employment rate, but this fell sharply until reaching a plateau between ages 30-40, at approximately 75 percent. At the same time, these younger cohorts have delayed births, but their completed fertility rate has increased. I build a life-cycle model of household labor supply and fertility decisions. Women accumulate on-the-job experience and face an increased risk of infertility with age. I calibrate the model to match the life-cycle profile of employment, hours worked, and fertility decisions of women born between 1944-1957. Supported by my empirical findings, I assume that two changes in the economic environment trigger the shift in employment and fertility decisions of younger college-educated married women: higher returns to experience, especially at younger ages, and the availability of infertility treatments. I find that on-the-job accumulation of experience explains the higher employment rate at younger ages and the delay in fertility. However, without infertility treatments, the model fails to produce the increase in the total fertility rate observed in the data. As a result, my paper demonstrates the importance of considering both drivers when comprehending young college-educated married women's employment and fertility decisions.

The second chapter *Cash Transfers and Fertility: From Short to Long Run* is a coauthored project with Francisco Javier Rodríguez-Román. This paper motivates by the fact that many developed countries are at risk of experiencing population

decline due to low fertility rates, with potential adverse economic effects. As a response, governments are deploying family policies to increase the number of children. In this paper, we propose a dynamic life-cycle model of fertility and female labor force participation to assess their effectiveness. We use the short-run fertility effects of a cash transfer policy implemented in Spain in 2007-2010 to calibrate its parameters. Using the calibrated model, we find that the impacts, in the long run, are half as large as in the short run. This is driven by differences in the responses of younger and older women at the time of implementation. The latter must react shortly after, as they cannot delay fertility much longer. The former anticipates their first birth. This generates additional births in the short run. We also study the effects of an alternative policy consisting of childcare subsidization and explore how the coexistence of temporary and permanent contracts in Spain, which have different earnings profiles, affects fertility and interacts with cash transfers by raising the costs of career interruptions in crucial child-bearing years.

In the third chapter, *Gender Gaps in the Labor Market and Social Security Finances*, I explore to what extent the incorporation of Spanish-married women into the labor market helps alleviate the aging process for Social Security finances. Although the increase in female labor force participation has been well documented and explained in the literature, no study has shown its impact on Social Security funding. To this end, I build a rich overlapping generations model with heterogeneous married households. The economy is initially on a balanced growth path in 1975 when several demographic changes trigger a transitional path, and eventually, the economy reaches a new balanced growth situation. My model matches well with the labor force participation of married women by educational attainment, hours worked, and key pension moments. It also generates an accurate life-cycle profile for savings. This paper shows that until 2050 there is a “women bonus”. Women finance around 10% of male pensions. Despite this, the model converges to a final steady state where gender differences persist in employment, hours worked, average earnings, and average pensions. In light of those, I evaluate the introduction of Gender-Based Taxation in the Spanish pension system. I assume that women are taxed at a lower permanent rate than men. This policy significantly narrows the gender gaps mentioned above and provides welfare gains for newborn cohorts.

The fourth chapter, *Female Portfolio Choice and Marital Property Regime*, is a coauthored project with Isabel Micó-Millán and Susana Párraga. This paper studies the link between married couples’ portfolio choices and property division rules. Using household data from the Spanish Survey of Household Finances, we exploit the regional variation in default marital property regimes in Spain to estimate the causal effect of property division rules on household financial investment. We find that separate-property couples hold riskier financial portfolios than community-property ones when wives are responsible for household finances. To rationalize this gap in risky asset holdings, we develop a financial portfolio choice model where wives

make savings decisions and couples differ in their property division rule. Divorce risk encourages higher precautionary savings in safe assets for community-property spouses compared to separate property due to higher dissolution costs of marital savings. This translates into separate-property spouses saving less and allocating a larger portfolio share to risky assets. Lower income levels and higher income risk for women reinforce this mechanism, contributing to explaining the property regime gap in risky financial investment between couples.

Contents

List of Figures

List of Tables

Chapter 1

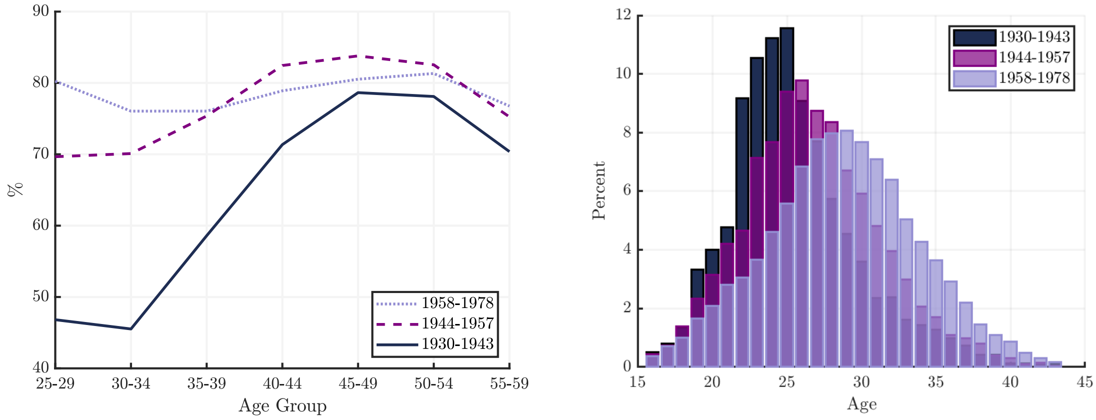
A Quantitative Theory of the New Life Cycle of Women's Employment

1.1 Introduction

In the United States, recent cohorts of college-educated married women have experienced dramatic changes in their employment profiles and fertility decisions, as shown in Figure 1.1. The employment profile of cohorts born before the fifties was initially flat and rose steeply after age 30. In contrast, the employment rate of women born after the mid-1950s started at a much higher level and fell sharply until reaching a plateau between ages 30-40, at approximately 75 percent (see Figure 1.1a). At the same time, younger cohorts have delayed births, but their completed fertility rate has increased (see Figure 1.1b). These observations motivate the following questions: why do college-educated women delay childbirth and work so much when they are young? Why does the employment rate remain low after age 30 and then flatten for a decade? How can these women have more children if they become mothers later in life? Why don't we see these changes in behavior among non-college-educated married women?

In this paper, I develop a quantitative theory to provide a unified explanation for the changes across cohorts of employment and fertility decisions of college-educated women. I build a life-cycle model of labor supply and fertility decisions. Children are costly in goods and mother's time. Women face an increasing risk of being infertile as they age. In turn, this induces women to become mothers early in their lives, reducing their labor supply until their children are grown and less expensive. In the model, labor market experience implies higher wages in the future, and the returns to experience are higher for younger women. Returns to experience increase the opportunity cost of being a young mother and induce highly productive women

Figure 1.1 Employment Age-Profile of Women and Maternal Age at First Birth



(a) Employment Rates (b) Mother's Age at First Birth

Source: March CPS-ASEC microdata (1963-2019) and CPS June Fertility Supplement (1976-2018). Note: The sample includes ever-married women with at least a college education. For the mother's age at first birth, I restricted the sample to women older than 16.

to work hard when young and to postpone births. In my model, two changes in the economic environment trigger the shift in employment and fertility decisions of younger college-educated married women. First, the increase in returns to experience, especially at younger ages, encourages these women to work more at the beginning of their lives and postpone having children, which can decrease their total number of children. Second, the availability of infertility treatments for younger cohorts lowers the opportunity cost of delaying fertility because of the increased chances of becoming a mother as they age. I find that accumulation of experience explains the higher employment rate at younger ages and the delay in fertility. Despite this, in the absence of infertility treatments, the model does not generate the increase in the total fertility rate we observe in the data. It is, therefore, important to consider both drivers when analyzing the employment and fertility decisions of young college-educated married women.

The first part of my analysis is empirical. First, I document and compare the life-cycle profiles of employment and fertility decisions for three cohorts of college-educated married women using the Annual Social and Economic supplement to the Current Population Survey (CPS-ASEC) microdata (1963-2019). The first includes those born between 1930-43, the second between 1944-57, and the third from 1958-78. I refer to these cohorts as Old, Middle, and Young, respectively. Women in the Young cohort appear to be planning a lifelong career. Compared to women in the Middle cohort, their employment rate between ages 25-29 is 14.3% higher, and at the same time, they delay children by three years. This causes a dip in labor supply between ages 30-39 (see Figure 1.1). Goldin and Mitchell (2017) named this drop in employment at age 30 the "sagging-middle". Despite delaying fertility, patterns in the data suggest that the total fertility rate of women in the Young cohort increased by 4.2% in comparison with women in the Middle cohort.[1]

Second, I show reduced-form evidence of a correlation between the sagging middle effect and fertility decisions. Using the CPS-ASEC microdata (1963-2019), I estimate the probability of being employed for college-educated married women. While the probability of being employed at ages 30-34 and 35-39 is 3.56 and 2.89 percent points smaller, respectively, than those at ages 25-29, once I introduce fertility controls in the regression, the probability becomes positive and statistically significant. These results imply that children are responsible for the decrease in employment among women in their thirties.

Third, I provide evidence of two exogenous explanations leading to the sagging middle effect and the fertility changes of the Young cohort. On the one hand, using the Panel Study of Income Dynamics (PSID) data, I find that college-educated married women born between 1958-1978 have higher returns to experience, especially at younger ages. In particular, returns to experience increased by 46%, relative to women born between 1944-57. Thus, there is an economic incentive to postpone fertility. On the other hand, there is an increase in the probability of women having children when old, thanks to Assisted Reproductive Technology (ART).[2] In the United States, some states cover or offer infertility treatments while others do not. The empirical has exploited this state-variation, and they find that the mandate significantly increases first birth rates for women over 35 and it increases the probability that women delay fertility (Schmidt, 2007; Buckles, 2007; Machado and Sanz-de Galdeano, 2015). Moreover, it is correlated with increased labor force participation for women ages 25-34 and decreased participation for women ages 35-44 (Buckles, 2007). Finally, they affect disproportionally older and highly-educated women (Bitler and Schmidt, 2012). These findings suggest that ART might have influenced career and family trade-offs in favor of postponing births, increasing fertility rates, and higher employment rates earlier in life.

This evidence suggests a link between the increase in returns to experience and the availability of infertility treatments with the employment and fertility decisions of Young cohorts. In the second part of my paper, I develop a quantitative life-cycle model of married individuals. Women choose their labor supply, fertility, consumption, and accumulate on-the-job experience. The husband's earnings have a deterministic age profile and a stochastic component. Children are costly in goods and in the mothers' time. Households can substitute only partially mothers' time with market childcare when mothers work outside the home. Each period, women face a fertility probability that decreases with age.

I calibrate the model to match the life-cycle profile of employment, hours worked,

[1] This is why Goldin (2021) says that these women "have it all: career and family".

[2] ART includes all fertility treatments in which either eggs or embryos are handled outside a woman's body.

and fertility decisions of women born between 1944-1957, the Middle cohort. The model replicates well the main features of the data. It generates an inverted U shape in employment, the average age at first birth, and the distribution of households by the number of children. I run two experiments to quantify the extent to which the new employment and fertility trends can be explained by gains in returns to experience and ART. In the first experiment, I consistently increase the returns to experience in the model with the returns to experience that I estimate using PSID data of the Young cohort. As a result, women increase their employment rate by 12% between ages 25-29 relative to the baseline economy, which is 0.5 percentage points higher than in the data. Moreover, these women delay having their first child by 1.38 years. Their first child is born at an average age of 28.8, which is similar to the average age of 28.9 in the data. Yet the model predicts that the total fertility rate will drop by 16.7%, whereas the data show an increase of 4%. This motivates the second experiment, where I increase the returns to experience and introduce ART. To capture the increase in the total fertility rate observed in the data, I calibrate the increase in the fertility probabilities after age 30 to match it. In this case, women postpone births, but their overall employment rate increases by less than in the first experiment. This is partly because more women have two and three children, which decreases employment relative to the results in the first experiment. Overall, the employment rate is largely in line with the data of the Young cohort in this experiment. My results suggest that an increase in returns to experience, especially at younger ages, can qualitatively account for both the increase in employment before age 30, the flat employment profile between ages 30 and 39, and the delay in fertility. Although on-the-job accumulation of experience plays a crucial role, the model does not generate an increase in the total fertility rate in the absence of infertility treatments. Therefore, it is essential to combine both factors to explain the new shape of the age profile of employment and the changes in fertility decisions of the Young cohort.

The contributions I make are both empirical and theoretical. My empirical contributions are two. I contribute to the main paper that illustrates the new employment life-cycle profile of college-educated married women, Goldin and Mitchell (2017), by providing reduced-form evidence of a correlation between the new employment profile and fertility decisions. Furthermore, I have also documented a new fact: these women are changing not only their employment profiles but also their work hours throughout their lives. There is a sagging middle effect in employment and hours worked. Second, I contribute to the literature on returns to experience by estimating them by cohort groups and by focusing on college-educated married women. Theoretically, this paper is the first attempt to link increasing returns to experience with fertility decisions under conditions of infertility treatment availability. Studies analyzing the impact of higher returns to experience typically focus on labor market outcomes rather than modeling fertility (Olivetti, 2006; Attanasio et al., 2008). Thus, this is the first paper to address the effects of higher returns to experience on fertility outcomes, such as the delay in births and the increase in the total fertility rate.

The rest of the paper is organized as follows. In Section 2, I position my paper in the context of the existing literature. In Section 3, I document the main changes in the labor market across cohorts in the U.S. In Section 4, I propose a set of potential explanations behind the change in women's behavior. In Section 5, I introduce the model. In Section 6, I carefully specify the calibration methodology. In Section 7, I analyze the calibration results. In Section 8, I disentangle the effect of each factor in shaping employment and fertility decisions. In Section 9, I quantify the effects of introducing individual taxation. In Section 10, I discuss why the proposed explanations have minor impacts on non-college-educated women in the U.S. Finally, in Section 11, I conclude.

1.2 Literature Review

The paper contributes to different strands of the literature. First, it contributes to the literature on the drivers of changes in female labor force participation, employment, and work hours over time in the US economy. A vast majority of papers have analyzed the causes behind the sharp increase in the female employment rate. These explanations include the power of the contraceptive pills (Goldin and Katz, 2002), the electricity revolution (Greenwood et al., 2005), the relative change in returns to experience compared with the men's (Olivetti, 2006), the decrease in the gender wage gap (Jones et al., 2015), the infant formula (Albanesi and Olivetti, 2016) and the reduction in the child care cost (Sánchez-Marcos and Bethencourt, 2018), among many others.

Despite the literature having well-analyzed the female labor market increase over time, less research has been done on the new employment profile of recent cohorts of women in the US. To the best of my knowledge, Goldin and Mitchell (2017) are the first to document a "sagging middle" in the labor force participation of American college-educated women. According to their study, younger cohorts accumulate more work experience than previous cohorts, especially at a younger age. In addition, they show that a significantly higher proportion of these women work at least 80 percent of the time between the ages of 25 and 54. In addition, they conduct event studies in which they determine that, unlike previous cohorts, college-educated women's labor force participation does not fully recover ten years after the birth of their first child. Finally, they demonstrate that an increase in the age at which the first child is born is correlated with increased participation during the 25-34 year interval, but decreased participation during the 35-44 year interval. Despite this, their analysis does not demonstrate a correlation between the sagging middle effect in employment and women's fertility choices.[3] Moreover, their analysis does not address the underlying factors responsible for the observed sagging middle in employment and fertility trends together. In addition, their study focuses on the extensive margin, whereas I also provide evidence of a sagging middle effect in the intensive margin. In this line, Buttet and Schoonbroodt (2013) build a life-cycle model of endogenous female labor supply to quantitatively explain why the employment profile of married women in the US born between 1940 and 1960 is flatter. They exogenously evaluate the role of the decrease and delay in births, the increase in relative wages of women to men, and the decline in childcare costs. They find that the relative wages explain 67% of the flatter life-cycle profile of employment.

Second, I also contribute to the literature examining possible causes behind aggregate labor force participation leveling off in the late 1990s. They range from changes in women's beliefs about the long-run payoff from working (Fernández, 2013), the convergence of information across regions (Fogli and Veldkamp, 2011), a retreat in egalitarian gender role attitudes, and the rebound in traditional gender role attitudes (Fortin, 2015), the lack of "family-friendly" policies (Blau and Kahn, 2013), a crowding-out effect in the labor market as they were replaced by college-educated men who used to work in blue-collar occupations (Duran-Franch, 2021) or growth in earnings inequality for women married to highly educated and high-income husbands (Albanesi and Prados, 2022). I contribute to this literature from a different angle. In particular, I introduce a life-cycle model to evaluate the extent to which changes in returns to experience and infertility treatments explain the change in the women's participation rate over the life cycle for the recent cohorts. Understanding the incentives women face to work over the life cycle, especially its recent shifts, is crucial to comprehending the overall stagnation.

Finally, this paper also contributes to the literature building structural life-cycle models of labor supply decisions of married individuals and fertility decisions to understand the interaction between them. Key contributions to this literature are Moffitt (1984), Caucutt et al. (2002), Francesconi (2002), Erosa et al. (2002), Da Rocha and Fuster (2006), Sheran (2007), among others. While fertility decisions play an important role in these papers, their primary focus is on labor market outcomes, returns to experience, and understanding the gender wage gap. My paper contributes to the most recent papers modeling fertility and labor supply decisions where the conflict between career-family is explicitly introduced. In this vein, Adda et al. (2017) estimate a dynamic life-cycle model of labor supply, fertility, and savings, incorporating occupational choices with specific wage paths and skill atrophy that vary over the career. They quantify the career cost of children when women drop out of the labor market due to childcare. In contrast, my focus is on understanding the causes of the postponement of births and the decline in employment between 30 and

[3]This is because they only consider households with children in this exercise

39 years of age. Eckstein et al. (2019) also estimate a life-cycle model of individual and household decisions regarding education, employment, marriage, divorce, and fertility. Similar to them, I show the relevant role of birth technology on career and family choices—they model contraception technologies, whereas I consider infertility treatments.

1.3 Facts to Explain

In this section, I provide empirical evidence about three facts. First, college-educated married women born after 1958 have a different life-cycle employment profile. This profile shows a sagging between the ages of 30-39. Further, I show that these women delay births compared to previous cohorts and have more children than those born in the 1940s-mid-1950s. Second, I demonstrate that the decrease in employment between ages 30-39 correlates with changes in fertility decisions. Third, I show that this new employment profile is a particular feature of college-educated married women.

1.3.1 New Life-cycle Profile of Employment

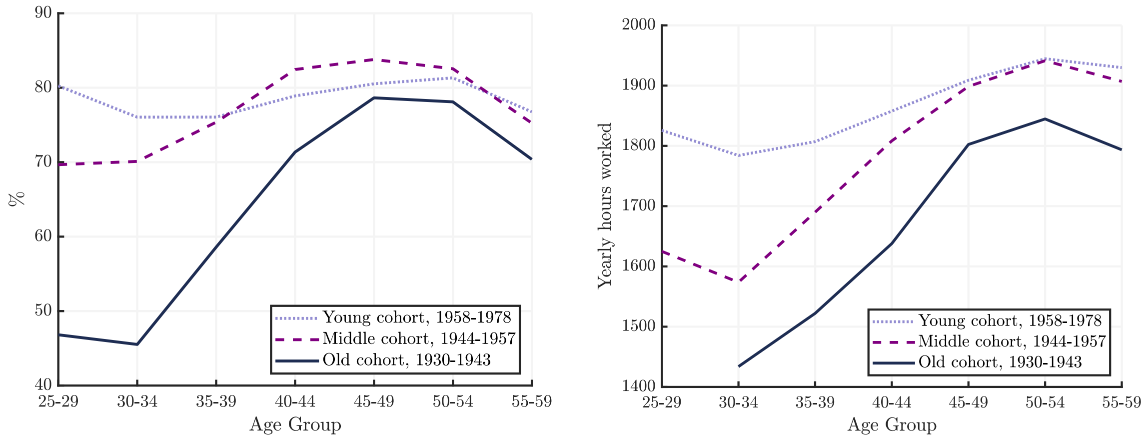
Data sources and sample. The primary datasets are the Current Population Census March, and the June Fertility Supplements extracted from the IPUMS for 1962-2019 and 1976-2018, respectively. I construct synthetic birth cohorts following Goldin and Mitchell (2017).[4] In particular, I restrict attention to three cohorts: the first includes those born between 1930-43, the second between 1944-57, and the third between 1958-78. I refer to these cohorts as Old, Middle, and Young, respectively. The sample consists of college-educated married individuals aged between 25-54.

Results. Figure 1.2 shows the employment rate and weekly hours worked for college-educated married women who belong to three different cohorts: Old, Middle, and Young.

The Old cohort includes women born between 1930-43. Most of these women become mothers in their twenties and have, on average, 2.5 children. Their employment rate is 48% between ages 25-30. After childbearing, these women increase their labor market attachment and reach an employment rate of almost 80% at the age of 45-49. Therefore, the average woman in the Old group has a *family and then a job* (Goldin, 2021).

[4] Synthetic birth cohort is similar to cohort analysis, but instead of using successive observations of the same group of people, you treat the population's age distribution as if it were a cohort passing through time.

Figure 1.2 Life-cycle Profile of Female Employment and Hours Worked



(a) Employment Rates (b) Yearly Working Hours

Source: CPS-ASEC microdata, March, 1963 to 2019.

Note: The sample includes ever-married women aged 25-59 with at least a college education. The variable yearly working hours is computed as the product of last year's weeks and the usual hours worked per week last year.

The Middle cohort includes women born between 1944-57. Unlike the Old cohort, they decide to get on the career track first at the expense of delaying births. These women have their first kid at the age of 26.7, and they have, on average, 1.9 children. Consequently, 27% of them never have a kid. Their employment rate is higher than the previous groups, but the life-cycle shape remains similar. They have a constant employment rate of 70% between the ages of 25-34 and an increase after that. In the 45-49 age group, it is 82%. Thus, the average woman in the Middle group has a *career and then a family* (Goldin, 2021).

Finally, the Young group includes women born between 1958-78. They continue delaying fertility. They have their first child at the age of 29. Despite this, on average, they have more children than the Middle cohort. Even though we cannot observe the completed fertility for all the cohorts included in this group, the average fertility rate increases to 1.98.[5] The employment rate is similar on average between the Middle and the Young cohorts; however, over the life cycle, it is not. Women in the Young cohort have a higher employment rate between ages 25-29, dropping from the labor market and reducing the hours worked between ages 30 and 39, compared to the Middle cohort. This decline in the labor market attachment has long-lasting effects. Their employment rate and hours worked do not differ from the Middle cohort after age forty. As a result, the average woman in the Young group has it all: *career and family* (Goldin, 2021). In what follows, I examine the link between the postponement in births, the fertility increase, and the drop in employment between

[5]See Doepke et al. (2022a) for a detailed review of the literature showing that the negative correlation between income and fertility has reversed.

ages 30-39 for college-educated married women in the Young cohort.

1.3.2 Evaluating the sagging-fertility relationship

In the previous section, I suggested a correlation between female employment sagging in the 30-39 age group and the delay and increase in fertility for the Young cohort. To identify the role of children in the decision to be employed between ages 30-39 compared to ages 25-29, I estimate the probability of being employed as follows:

$$Prob(E_{it}) = \beta_0 + \sum_{j=1}^{6} \beta_j \text{Agegrpj}_{it} + \beta_6 X_{it} + u_{it}, \tag{1.1}$$

where E_{it} is 1 if the individual i at time t is employed and 0 otherwise. I consider 6 age groups of 4-year age bins starting at age 25, Agegrpj.[6] The omitted age group in the regression is Agegrp1, corresponding to ages 25-29. In the regression, I introduce a set of individual-level control variables, X_{it}, including the state where each member of the household was born, the husband's worker characteristics and education, wave identifiers, fertility characteristics, and u_{it} is the remaining unobserved heterogeneity. I estimate equation (1.1) through a Logit regression and compute the marginal effects.

Variables. The dependent variable in the model is employment status, which indicates if the individual is employed or not. To construct this variable, I consider an individual employed if she/he is in the labor force and the employment status is "at work", "has a job, not at work last year", or "Armed forces". On the contrary, an individual is not employed if she/he is in the labor force as "Unemployed". I differentiate two groups of controls. First, demographic controls include the individual's age group, the individual's native indicator, equal to one if the individual was born in U.S. territory, an indicator for college education, and a dummy when children live in the household. Second, the employment husband's characteristics include an indicator for a full-time-full-year worker, an indicator for the worker being a private, public or self-employed employee, and the pre-tax wage and salary income in logs. I consider that the individual is full-time full-year if she/he answers that she/he was a full-time worker and worked at least 50 weeks the previous calendar year.

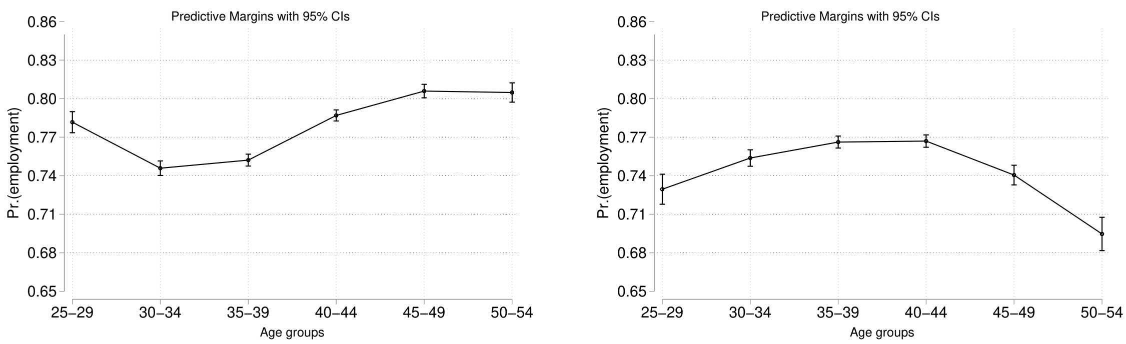
Methodology. To understand the role of fertility on employment decisions, I run the regression in equation (1.1) for four model specifications. Model (I) does not control by fertility decisions, while Models (II-IV) do. The differences between these last three are the set of controls I include. In the Appendix, Table A1.1 shows the

[6] Age groups from 1 to 6 correspond to the age intervals 25-29,30-34,35-39,40–44,45-49,50-54, respectively.

marginal effects of each control variable on the probability of being employed for those college-educated married women in the Young cohort group. The coefficients of interest are $\beta_2 - \beta_6$.[7] The interpretation of them is the following: if they are negative (positive), it means that the probability of being employed is lower (higher) in those age groups compared to the age group 25-29.

Results. The main result is the following: While the probability of being employed at ages 30-34 and 35-39 is 3.56 and 2.89 percent points smaller than those at ages 25-29 in Model (I), once I control for the presence of a child in the household, the number of children and the age of the youngest child, Model (IV), the probability becomes positive and statistically significant. In particular, the probability of being employed increases by 2.46 and 3.7 percent points for the 30-34 and 35-39 age groups compared to the 25-29. This implies that children are the main driving force behind the lower employment rate between ages 30-39. To illustrate this result, Figure 1.3 plots the predicted employment probability by age of Model (I) in Figure 1.3a and Model (IV) in Figure 1.3b. Only the first model predicts a sagging middle effect because there is no sagging middle effect once fertility decisions are considered.

Figure 1.3 Marginal Probability of Employment for College-Educated Women in the Young Cohort



(a) Not Fertility Controls

(b) Fertility Controls

Source: CPS-ASEC microdata, March, 1963 to 2019.

Note: The sample includes ever-married women with at least a college education. The regression is Table A1.1.

I confirm these findings by showing the employment and hours worked life-cycle profiles for two population subgroups. First, the shape is similar between the Middle and Young cohorts of women without children (see Figure A1.1 in the Appendix). Second, the sagging middle effect is more pronounced in households with more kids and a higher postponement in births. This is the case, for instance, of college-educated women married to college-educated men. The Young cohort has 0.13

[7] Table A1.2 and A1.3 show the results for Models (I) and (IV), respectively, for the Young, Old and Middle cohorts.

kids more, and they have them on average two years later compared to college-educated women married to non-college-educated men.[8] For the later group, their shape displays a sagging middle effect, but it recovers faster, and the dip is less pronounced (see Figure A1.2 in the Appendix).

1.3.3 Narrative of College-Educated Married Women

Despite providing evidence that college-educated married women have changed their employment profiles, I have not explained why this story is about married women with college degrees. Only those individuals exhibit a sagging middle effect in their employment profiles. To provide evidence for this, I compare three measures for labor market outcomes (employment, hours worked, and participation) for different sub-samples of the population. First, for non-college-educated married women, each line is above its predecessor, and there is no sagging middle effect (see Figure A1.3 in the Appendix). Second, for non-married-college-educated women, no clear patterns can be derived from this analysis (see Figure A1.4 in the Appendix). Third, for college-educated married men, the life cycle shows the same pattern across cohorts (see Figure A1.5 in the Appendix). Interestingly, a lower fraction of these men from the Young cohort are employed; however, this new phenomenon is out of the scope of the paper.[9]

Lastly, as a robustness check, I run the same regressions in Equation 1.1 for married women who are not college graduates. The sagging middle effect is not present for these women, even in Model (I). Figure A1.6 shows their predicted employment by age.

1.4 Potential Drivers

In this section, I provide detailed empirical evidence to support my theory. First, I demonstrate that college-educated women born after the mid-1950s experience greater returns to experience, particularly at younger ages. Second, I present evidence from the empirical literature suggesting that Assisted Reproductive Technology influenced women's decisions regarding employment and fertility in a way consistent with the behavior of women in the Young cohort.

[8] Table A1.5 displays the evolution in the average number of children in the household and the average age at which college-educated women have their first kid through the education of their husbands.

[9] See Juhn and Potter (2006), Krueger (2017) or Abraham and Kearney (2020) for an explanation of the decrease in employment to population ratio in the U.S.

1.4.1 Returns to experience

My analysis relies on data from the Panel Survey of Income Dynamics (PSID), a longitudinal survey of a representative sample of the American population. The University of Michigan runs the PSID, which has been conducted annually since 1968 and biennially since 1997. For my analysis, I used data from 1968 to 2015. My main argument for higher returns to experience for the Young cohort compared to women in the Middle cohort is based on three pieces of evidence: full-time workers' earnings have a steeper age profile, experience provides greater marginal returns, and higher returns to experience are associated with greater returns at a young age. Finally, I discuss and compare my findings with those of the literature.

Steeper wage profile of earnings for full-time workers. I use the panel structure of the PSID to construct a dataset where I restrict attention to married women for whom I have observed their labor market outcomes (at least) since age 25. I divide these women into two categories: almost full-time workers and all-sample workers. The first includes women who work full-time at least 80% of the time they are on the panel.[10] The second does not make this restriction; therefore, it includes all women who work regardless of their working hours.

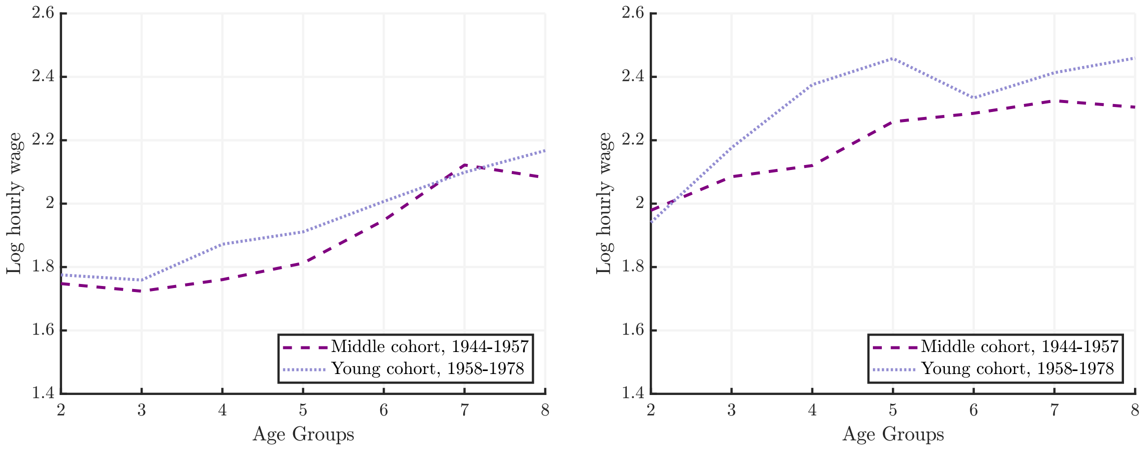
Figure 1.4 plots the median real log-hourly wage for all sample and full-time married women workers as a function of age for the Middle and the Young cohorts. A steeper wage profile suggests larger returns to experience. The difference between the all-sample-workers (left panel) and almost full-time-workers (right panel) shows that the former has lower returns to experience than the latter. Moreover, this Figure also shows that women in the Young cohort have higher returns to experience than those in the Middle cohort. This conclusion comes from two features. First, the hourly wage difference between the all-sample workers and almost-full-time workers is higher for the Young cohort. Second, almost full-time workers in the Young cohort display a steeper earnings profile than the Middle cohort. This finding provides the first evidence of a higher effect of experience on earnings, especially for young individuals.

Steeper experience-wage profiles. The second piece of evidence shows differences between the two cohort groups regarding experience-wage profiles. To do so, I use the couples dataset and restrict attention to college-educated married workers. I construct two different variables for the experience. First, potential experience is defined as the number of years that have elapsed since a worker finished schooling or turned 18, whichever is smaller.[11] Second, actual experience is constructed with

[10] I consider that an individual is a full-time worker if she works at least 1500 hours.

[11] As in Lagakos et al. (2018), I define potential experience as experience = age - schooling - 6 for individuals with 12 or more years of schooling and as experience = age- 18 for individuals

Figure 1.4 Real Median Log Hourly Wages for Married Women by Hours Worked



(a) All-Sample-Workers (b) Almost Full-Time-Workers

Source: PSID, 1968 to 2015.

Note: Sample: married women. I use the 1968 family weight. Almost full-time workers imply that they work at least 80% of the time as full-time workers through the panel. I consider an individual a full-time worker if she works at least 1500 hours.

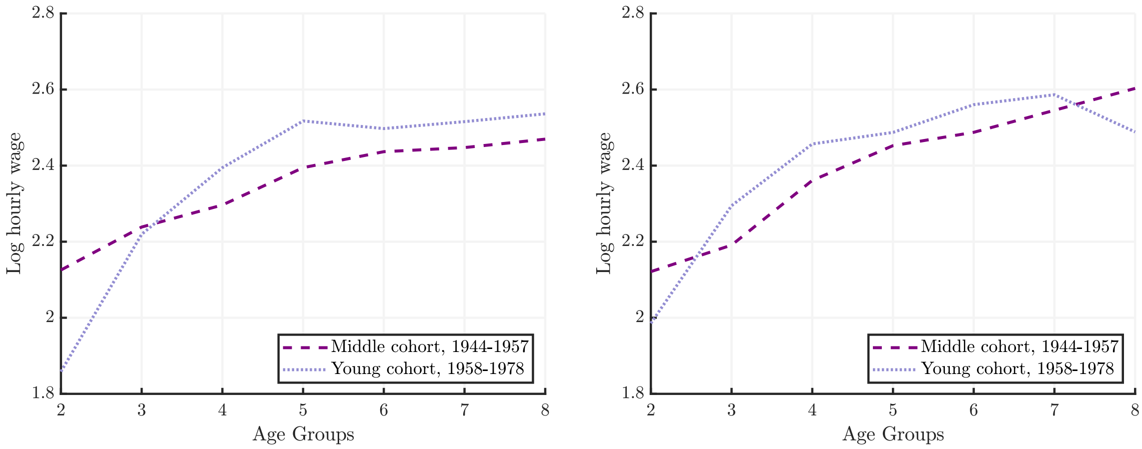
the family question about the number of years the head and the spouse have worked full-time since they were 18.

Figure 1.5 shows the experience-wage profiles for these two experience measures. The main conclusion is that women in the Young cohort experience a much steeper experience-wage profile. Specifically, the wage growth between the first two groups of experience is higher in the Young cohort than in the Middle one. Thus, it is more likely that those women delay fertility to accumulate enough experience compared to the Middle cohort, for whom the difference is smaller.

To complement the previous findings, I run a Mincer regression for each cohort group of college-educated married women. I regress log hourly wages on a constant, experience, and experience squared. Table 1.1 shows the coefficients of such regressions. According to this analysis, women in the Young cohort experience higher returns to experience than those in the Middle cohort. The coefficient for experience is 3 pp higher. Experience squared, however, also has a higher coefficient, suggesting that as women accumulate experience, they will have more similar returns to those of the Middle cohort. Figure 1.6 shows the average marginal effect of one extra year of experience for a given level of experience stock. After eight years of labor market experience, marginal returns of experience for the Young cohort become not statistically different from the Middle cohort. It is important to note that the sample of women who have worked more than 16 years for the Young cohort is smaller than for the Middle cohort. Overall, the average rate of return to a year of experience

with fewer than 12 years of schooling.

Figure 1.5 Real Median Log Hourly Wages for Married Women by Experience Measure



(a) Potential Experience (b) Actual Experience

Source: PSID, 1968 to 2015.
Note: The sample includes college-educated married women. I use the 1968 family weight.

among the Middle cohort is 2%, while it is 3.43% for the Young cohort, i.e., it is 70% higher for the Young cohort.

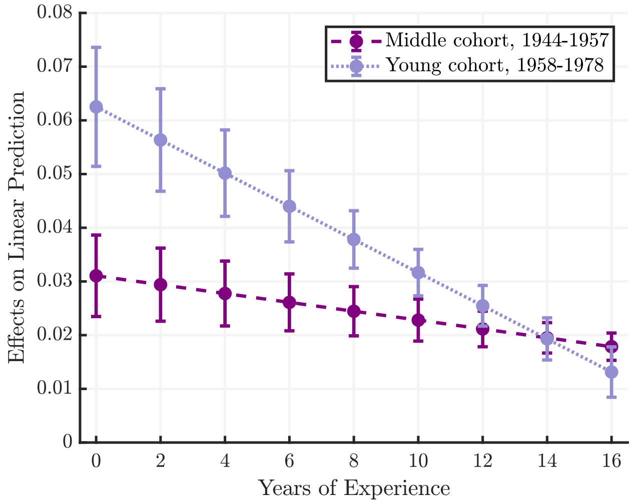
Table 1.1 Mincer Regression for College-Educated Married Women by Cohort (I)

	Middle Cohort	Young Cohort
Experience	0.0311***	0.0625***
	(0.004)	(0.006)
Experience^2	-0.000413***	-0.00154***
	(0.000)	(0.000)
Constant	1.974***	1.875***
	(0.030)	(0.030)
Observations	5278	3629
R-squared	0.0427	0.0640

Standard errors in parentheses
Source: PSID 1968-2015
Sample: College-educated married women
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Higher returns to experience when young. Finally, I demonstrate that the higher returns to experience are particularly higher at younger ages. First, I regress the daily log income on age, experience, and its interactions (treated as continuous variables). Table 1.2 shows the coefficients of these two regressions. From this table, it is clear that the coefficient of experience is greater for the Young cohort. Moreover, I plot in Figure 1.7 the average marginal effect of experience at different

Figure 1.6 Average Marginal Predicted Effect of Experience From Table 1.1



Source: PSID, 1968 to 2015.

Note: This Figure represents the conditional marginal effects of experience with 95% confidence intervals by cohort group.

ages. It indicates that the marginal returns to experience are greater for the Young cohort before age 30. Secondly, I run a Mincer regression where I control by an age dummy equal to one if the worker's age is younger than 30, experience, experience squared, and their interactions. Table 1.3 shows the results of this estimation. The coefficient of interest is the interaction between the age dummy and the experience. This shows that the marginal effect of experience on log-hourly wages is positive and significant for the Young cohort, while for the Middle one, there are no differences. This is suggestive evidence that Young cohorts are profiting from an increase in the returns to experience when young. The results of the last regression are robust to the inclusion of more control variables, including the number of children, the husband's job characteristics, experience, and education, the use of only one regression for women born in the Young cohort, and eliminating the squared term of experience. See Tables A1.6, A1.7 and A1.8 in the Appendix.

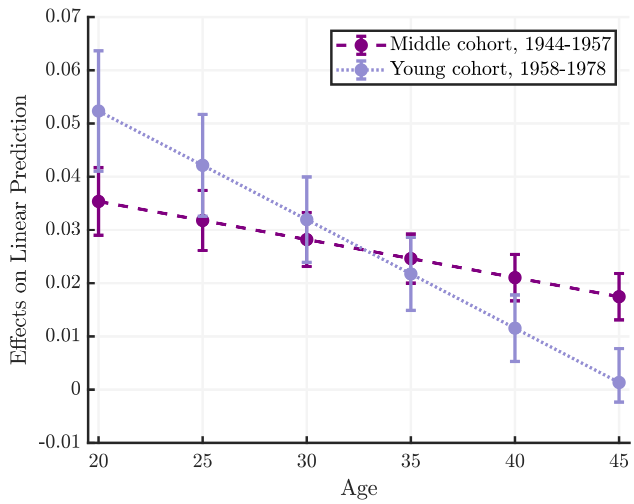
Discussion My hypothesis is that college-educated women born after the mid-1950s increase their employment rate at younger ages because of increased returns to experience. In addition, greater returns to experience at younger ages motivate women to delay fertility because children impose a time cost for mothers reducing their labor supply. This mechanism is in line with Olivetti (2006). She shows that between the 1970s and 1990s, women's returns to experience increased more than men's, which helps explain a significant portion of the increase in female employment. Nonetheless, she does not attempt to make the relationship between postponing births and increasing returns to experience. The positive effect of fertility delay on wages is in line with the empirical findings. Amuedo-Dorantes and Kimmel

Table 1.2 Mincer Regression for College-Educated Married Women by Cohort (II)

	Middle Cohort	Young Cohort
Age	0.0102***	0.0333***
	(0.002)	(0.003)
Experience	0.0497***	0.0931***
	(0.005)	(0.010)
Age × Experience	-0.000716***	-0.00204***
	(0.000)	(0.000)
Constant	1.639***	0.968***
	(0.078)	(0.089)
Observations	5420	3704
R-squared	0.0437	0.0809

Standard errors in parentheses
Source: PSID 1968-2015
Sample: College-educated married women
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.7 Average Marginal Predicted Effect of Experience From Table 1.2



Source: PSID, 1968 to 2015.
Note: This Figure represents the conditional marginal effects of experience with 95% confidence intervals by cohort group.

(2005) find that college-educated mothers experience a wage boost of about 4 percent compared to their non-college-educated counterparts. They point to selecting more family-friendly jobs for mothers as the leading explanation. Moreover, they find that college-educated women additionally enhance a wage boost by 13 percent when delaying fertility. This could suggest that women accrue the maximum benefit to their human capital investment by postponing fertility. Miller (2011) shows that

Table 1.3 Mincer Regression for College-Educated Married Women by Cohort (III)

	Middle Cohort	Young Cohort
Age dummy (< 30)	-0.0843	-0.529***
	(0.059)	(0.070)
Experience	0.0279***	0.0281***
	(0.005)	(0.008)
Age dummy (< 30) × Experience	-0.0124	0.135***
	(0.010)	(0.024)
Experience^2	-0.000391**	-0.000644*
	(0.000)	(0.000)
Age dummy (< 30) × Experience^2	0.000190	-0.0130***
	(0.001)	(0.002)
Constant	2.035***	2.171***
	(0.048)	(0.057)
Observations	5420	3704
R-squared	0.0433	0.0800

Standard errors in parentheses

Source: PSID 1968-2015

Sample: College-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

motherhood delay leads to a substantial increase in earnings of 9% per year of delay, an increase in wages of 3%, and an increase in work hours of 6%.

1.4.2 Assisted Reproductive Technology (ART)

Assisted Reproductive Technology (ART) is defined by the Centers for Disease Control and Prevention (CDC) as "all treatments or procedures that include the handling of human eggs or embryos to help a woman become pregnant". The main type of ART is *in vitro* fertilization (IVF), and the first successful procedure in the United States was in 1981. The availability of ART significantly changed women's reproductive lives, as it increased the chances of becoming a mother at older ages. However, like in many European countries, infertility treatments are relatively new, costly, and many insurance plans do not cover them. According to Hamilton and McManus (2012), each attempt of IVF treatment typically entails a cost of $10,000 to $15,000 to the patient, with a probability of success around 25-30%.[12]

Despite this, in the United States, 17 states have passed legislation requiring

[12]See Bitler and Schmidt (2012) for an analysis of the cost of different infertility services.

insurance companies to cover varying degrees of infertility treatment since the 1980s, and a few states mandate IVF coverage specifically. In some states, however, this coverage is restrictive and not available to everyone.[13] Hamilton and McManus (2012) showed that insurance mandates for IVF increased both treatment access and reduced the aggressiveness of the procedure.

The empirical literature has exploited the variation in ART coverage across states to show how infertility treatments affect women's employment and fertility decisions, among other outcomes. Buckles (2007) shows that the presence of this law is correlated with increased labor force participation for women ages 25-34 and decreased participation for women ages 35-44. Buckles (2007) and Machado and Sanz-de Galdeano (2015) demonstrate that the coverage of these services increased the probability that women delay fertility. Schmidt (2007) shows that these mandates significantly increase first birth rates for women over 35. Moreover, as discussed in Moreno-Maldonado and Santamaria (2022), the mandate's effect on the delay of parenthood went above the increase in the utilization rates. ART can incentivize women to postpone fertility, even if they don't use them in the end. Additionally, it changed women's perceptions of its effectiveness.

ART availability generates similar employment and fertility decisions to those of married women with a college degree in the Young cohort. Since this technology is recent, it can only affect the employment and fertility decisions of women in the Young cohort.[14] Further evidence of this is the fact that these mandates disproportionately affect older and college-educated women (Bitler and Schmidt, 2012). For instance, Abramowitz (2017) employs duration and competing risks analyses, exploiting the exogeneity of mandates, to document that mandates are associated with delayed marriage and childbearing at younger ages and an increased likelihood of marriage and motherhood at ages 30 and older, but only for college graduate women.

In a nutshell, in this section, I have shown that the employment decrease between ages 30-40 for college-educated married women in the US correlates with the delay in births and the increase in the number of children. Two potential drivers might explain these employment and fertility trends. The increase in the returns to experience, especially at younger ages, might have created an economic incentive for women to work more when in their twenties and postpone births. Besides, the importance of higher returns to experience at younger ages is accentuated by the fact that assisted reproductive technologies have decreased the incidence of infertility among women as they age. Taking these factors into account, I develop the

[13] Arkansas, Connecticut, Delaware, Hawaii, Illinois, Louisiana, Maryland, Massachusetts, Montana, New Hampshire, New Jersey, New York, Ohio, Rhode Island, and West Virginia are among the states that require insurers to cover diagnosis and treatment for infertility. However, of those 17 states, two states require insurers to offer coverage for infertility diagnosis and treatment: California and Texas.

[14] Women in the Middle cohort were years old at least when ART treatments were available and

life-cycle model I introduce in the following section.

1.5 The Life-Cycle Model

This section describes the model I develop to explain the new life-cycle profile of women's labor supply. The economic environment consists of a discrete-time life-cycle model of married individuals who never divorce. Women make labor supply, fertility, and consumption decisions. There is on-the-job experience accumulation. The model reflects women's main trade-offs when making career and fertility decisions. There are monetary and time costs associated with becoming a mother. In addition to childcare services, women also suffer wage losses due to interruptions in their labor market participation. The model hinders women's earnings growth because there is no experience accumulation if they do not participate in the labor market.

1.5.1 Demographics

Each household is composed of a wife and a husband, $g = \{f, m\}$, who never divorce. I assume that both spouses are of the same age and face the same deterministic lifespan. They enter the economy at the age of $j = 1$, corresponding to age 25, and they die at age $j = J$, which is age 55 in reality. As one model period is one year, $J = 31$.

Between ages $j \in [1, \bar{J}]$, the wife decides whether she wants to have an additional child. I denote this decision by $b_{j+1} \in \{0, 1\}$. Each woman can have a maximum of three children, and I denote the total number of children by $N_j \in \{0, 1, 2, 3\}$. Regardless of her decision ($b_{j+1} = 1$), she gives birth with probability $p^j(N_j)$. This probability is a function of age to capture increasing infertility with age. It is also a function of the number of children already in the household.

After being born, children age stochastically.[15] They belong to four different groups: newborns n_0, babies n_1, school-age children n_2 or teenagers n_3. These stages differ in the cost children impose on mothers as they age and the utility they derive from them. Newborns and babies entail the same consumption, leisure, and childcare costs. However, newborns provide extra utility for mothers who stay at home. School-age children continue to impose consumption and leisure costs, although quantitatively not identical to those imposed by newborns and babies. Moreover, I assume that there is a universal public provision of school, and thus, school-age children do not impose childcare costs. Finally, those in the fourth stage, teenagers/young adults, are not costly. The exact mathematical form all these costs

[15] Da Rocha and Fuster (2006) and Guner et al. (2021) also assume stochastic aging of children to reduce the number of state variables.

take and the functional form of the utility of children are discussed in subsections 1.5.3 and 1.5.6.

I denote by $\mathbf{n_j} = \{n_{0j}, n_{1j}, n_{2j}, n_{3j}\}$ the vector indicating the number of children of each type, and hence $N_j = n_{0j} + n_{1j} + n_{2j} + n_{3j}$ is the total number of children in the household of age j. A newborn becomes a baby with a probability equal to one next period. A baby becomes a school-age child, and a school-age child becomes a teen with probabilities λ_1 and λ_2, respectively. The teenage/young adult stage is absorbing: once a child reaches this stage, it remains there. I denote this structure by $\mathbf{n_{j+1}} = \Lambda(\mathbf{n_j}, b_{j+1})$[16].

1.5.2 Time Allocation

Each individual is endowed with one unit of time. I assume that husbands always work full-time. On the contrary, women make fertility, market work, and consumption decisions. Thus, women are the sole decision-makers in the household.

Each period, besides their fertility decision, women decide their labor supply decision (h_j^f). I model the extensive and intensive margins of market work. Thus women decide between participating or not in the labor market, and if they do so, they decide between becoming part-time or full-time workers. I refer to these three labor market statuses as $h_j^f \in \{0, pt, ft\}$, respectively.

1.5.3 The Cost of Children

Children are costly in the mother's time. The cost increases with the number of kids but decreases with the child's age. In particular, the time cost associated with having children is:

$$\tau_j^f(\mathbf{n_j}) = \begin{cases} \zeta_1(n_{0j} + n_{1j}) & \text{if } (n_{0j} + n_{1j}) > 0 \\ \zeta_2 n_{2j} & \text{if } (n_{0j} + n_{1j}) = 0 \text{ and } n_{2j} > 0 \\ 0 & \text{if } (n_{0j} + n_{1j} + n_{2j}) = 0 \end{cases}, \tag{1.2}$$

where $\zeta_1 > \zeta_2$. As a result of this functional form, a woman who is taking care of newborns and babies is also able to take care of school-aged children without incurring any additional time.

In addition to time costs, children are costly in terms of goods. First, mothers who work need to buy childcare services in the market for newborns, babies, and school-age children. These childcare services are proportional to the labor status of the mother. In particular, I assume that the following function describes the childcare services:

[16] See appendix A1.3 for the exact functional form this structure takes.

$$s_j(\mathbf{n_j}) = \begin{cases} \left(\kappa_1 \left(n_{0j} + n_{1j}\right) + \kappa_2 \left(n_{2j}\right)\right) \bar{Y}_j & \text{if } h_j^f = ft \\ 0.5 \left(\kappa_1 \left(n_{0j} + n_{1j}\right) + \kappa_2 \left(n_{2j}\right)\right) \bar{Y}_j & \text{if } \mathrm{h}_j^f = pt \end{cases}, \tag{1.3}$$

where $\bar{Y}_j$ denotes the average household income. Second, children are costly because they deflate consumption, $\Psi(\widetilde{N}_j)$. I denote by $\widetilde{N}_j = n_{0j} + n_{1j} + n_{2j}$ the number of costly children in the household at age j. Note that young adults do not incur time or goods costs.

1.5.4 Experience and Earnings Dynamics

Household income is the sum of the husband's and wife's labor earnings. Earnings of each individual of gender g at period j, is the product of labor productivity z_j^g and the hours worked in the market h_j^g :

$$w_j^g = z_j^g h_j^g. \tag{1.4}$$

The husband's labor productivity, z_j^m, is a function of age because I assume they always work full-time. However, females' productivity, z_j^f, depends on their experience accumulation at age j, x_j^f. The wage processes are:

$$\begin{aligned} \ln\left(z_j^m\right) &= \eta_0^m + \eta_1^m j + \eta_2^m j^2 + u_i^m + v_j^m, \\ \ln\left(z_j^f\right) &= \eta_0^f + \eta_1^f x_j^f + \eta_2^f \left(x_j^f\right)^2 + u_i^f + v_j^f, \end{aligned} \tag{1.5}$$

where $u_i^g \sim N\left(-\frac{\sigma_{u^g}^2}{2}, \sigma_{u^g}^2\right)$ is a fixed effect at birth, and v_j^g is a persistent productivity shock. Both wives and husbands receive uncorrelated shocks v_j^f and v_j^m, which evolve stochastically over time according to an AR(1) process:

$$\begin{aligned} v_j^m &= \rho v_{j-1}^m + \xi_j^m \\ v_j^f &= \rho v_{j-1}^f + \xi_j^f \end{aligned} \quad \text{with} \begin{bmatrix} \xi_j^m \\ \xi_j^f \end{bmatrix} \sim N\left(\begin{matrix} -\frac{\sigma_{\xi^m}^2}{2} \\ -\frac{\sigma_{\xi^f}^2}{2} \end{matrix}, \begin{bmatrix} \sigma_{\xi^m}^2 & 0 \\ 0 & \sigma_{\xi^f}^2 \end{bmatrix} \right), \; v_0^m = v_0^f = 0. \tag{1.6}$$

I assume that when the wife works full-time, she accumulates one year of experience. However, if she works part-time, she accumulates half-year experience. She keeps the same experience if she does not work. Thus, female experience in the labor market accumulates as follows:

$$x_{j+1}^f = F(x_j^f, h_j^f) = \begin{cases} x_j^f + 1 & \text{if } h_j^f = ft \\ x_j^f + 0.5 & \text{if } h_j^f = pt \\ x_j^f & \text{if } h_j^f = 0 \end{cases}. \tag{1.7}$$

1.5.5 The Government

The government collects taxes from labor income, and the tax's revenue is simply wasted. The tax function is

$$T(Y_j) = \left(1 - \lambda Y_j^{-\tau}\right) Y_j \tag{1.8}$$

where τ measures the degree of progressivity, λ is the average level of taxation and Y_j is the total household income. This functional form allows for negative tax rates, and thus it incorporates the Earned Income Tax Credit (EITC).

1.5.6 Preferences

Each period the household receives utility from consumption (c_j), leisure (l_j^f), and children ($\mathbf{n_j}$). Each child deflates consumption, $\Psi(\widetilde{N}_j)$, and implies a time cost that reduces leisure, $\tau_j^f(\mathbf{n_j})$. In addition, mothers derive utility from the number of costly children they have, $\widetilde{N}_j$, and the value of staying at home when the child is a newborn, υ. Based on this, we can formulate the per-period utility function as follows:

$$\begin{aligned} U_j(c_j, l_j^f, \mathbf{n_j}) =& \alpha_c \frac{\left(\frac{c_j}{\Psi(\widetilde{N}_j)}\right)^{1-\gamma_c} - 1}{1-\gamma_c} + \alpha_l \frac{(l_j^f)^{1-\gamma_l} - 1}{1-\gamma_l} + \alpha_n \frac{\left(1 + \widetilde{N}_j\right)^{1-\gamma_n} - 1}{1-\gamma_n} \\ & - \chi \mathbb{1}_{N_j > 0} + \upsilon \mathbb{1}_{n_{0j}=1} \mathbb{1}_{h_j^f = 0}, \end{aligned} \tag{1.9}$$

where χ is a fixed disutility of motherhood. This parameter captures either preference over not being a mother or other costs associated with motherhood that are not modeled. Anticipating the calibration, χ is crucial for the extensive margin of fertility (i.e. remaining childless or becoming a mother), while α_n and γ_n are important for the intensive margin (having 1, 2, or 3 children).

1.5.7 Budget and Time Constraints

The per-period budget constraint restricts the consumption in the household and expenditure on childcare activities to be equal to the sum of net household income. It is described by:

$$c_j + s_j(\mathbf{n_j}) = Y_j - T(Y_j), \tag{1.10}$$

where the household's income is the sum of both partners' labor income, i.e., $Y_j = w_j^f + w_j^m$. Regarding the time constraint, it restricts the sum of the time allocated to work, childbearing, and leisure to be equal to the endowment of time:

$$h_j^f + \tau_j^f(\mathbf{n_j}) + l_j^f = 1. \tag{1.11}$$

1.5.8 The Household Problem

I use a recursive formulation to describe the household's problem. The vector of state variables is given by the age, the experience accumulation of the wife, the vector of children by age, and the income shocks, $\mathbf{S} = \{x^f, \mathbf{n}, v^m, v^f\}$. Women decide how much to consume (c), how much to work (h^f), and whether to give birth or not next period (b'). The decision problem of a female at age j is given by:

$$
\begin{aligned}
& V_j\left(x^f, \mathbf{n}, v^m, v^f\right) = \\
& \max_{b' \in B(j,\mathbf{n}), c, h^f \in \{0, pt, ft\}} \left\{ U\left(c, \mathbf{n}, h^f\right) + \beta E\left[V_{j+1}\left(x^{f\prime}, \mathbf{n}', v^{m\prime}, v^{f\prime}\right) \middle| \mathbf{n}, v^m, v^f\right]\right\} \\
& \text{subject to} \\
& c + s(\mathbf{n}) = Y - T(Y) \\
& w^f = z^f h^f \\
& w^m = z^m h^m \\
& h^f + \tau^f(\mathbf{n}) + l^f = 1 \\
& x^{f\prime} = F(x^f, h^f) \\
& \mathbf{n}' = \Lambda_j(\mathbf{n}, b') \\
& \text{Equations 1.5, 1.6}
\end{aligned}
\tag{1.12}
$$

where the expectation is taken over the number of kids in each category and the productivity shocks. The choice set for the birth decision is defined as:

$$
B(j, \mathbf{n}) = \begin{cases} \{0, 1\} & \text{if } N < 3 \text{ and } j < \bar{J} \\ \{0\} & \text{otherwise.} \end{cases} \tag{1.13}
$$

1.6 Calibration

I calibrate the model using a two-step strategy to match the data for college-educated married women born between 1944 and 1957. In the first step, I use data to estimate the parameters that can be identified outside the model. Table 1.4 shows the value of these parameters. In the second step, I calibrate the remaining parameters to match the labor market life-cycle patterns and fertility decisions observed in the data. Although parameters and moments do not have a one-to-one mapping, I discuss how certain moments can provide information about some parameters.

1.6.1 First Step

Timing. The length of the model period corresponds to one calendar year. I only model two life stages. The first stage in life corresponds with the working periods when women are fertile. I assume that women begin their life at age 23, $j = 1$, and

Table 1.4 Calibration: Exogenous Parameters

Parameter	**Value**	**Meaning**	**Source**
Demographics			
J_f	21	Last fertile period	See text
J	37	Last model period	See text
Λ_1	0.2	Prob. from baby to school-age children	See text
Λ_2	0.1	Prob. from school-age children to teenager	See text
$p^j(0)$	-	Fertility probability by age if no children	See text
Utility			
ψ	$1.5 + 0.3(n_0 + n_1 + n_2)$	Equivalence scale	OECD
Income			
η_0^m	0.692		
η_1^m	0.078	Regression log wage on age and age2 (men)	
η_2^m	-0.001		
σ_{ξ^m}	0.173	St. dev. of husbands' income shock	
ρ^m	0.946	Persistance of husbands's income shock	PSID
σ_{ξ^f}	0.185	St. dev. of wives' income shock	
ρ^f	0.886	Persistance of wives' income shock	
σ_{u^m}	0.220	St. dev. of fixed effect for husbands	
σ_{u^f}	0.213	St. dev. of fixed effect for wives	
λ	2.915		Borella et al. (2022)
τ	0.107		Borella et al. (2022)

they are fertile until age 43, $j = \bar{J} = 21$. During this first stage, they face a fertility probability that decreases with the woman's age and the number of children.[17] The second stage corresponds with the periods $j \in [\bar{J}+1, J]$ when women are no longer fertile, but their children are still newborns, babies, or school-age children. The last period corresponds to the extreme case when the woman gives birth at the last fertile age, $j = \bar{J}$, and this child becomes a teenager, i.e., $J = 37$.[18]

I model four types of children according to their age: newborns (aged 0-1), babies (aged 1-5), school-age children (aged 5-14), and teenagers (+14). Children qualify as newborns for only one period, and they age deterministically, i.e. with probability one, a newborn at period j will be a baby at period $j+1$. However, transitioning from baby to school-age child and from school-age child to teenager is stochastic. The transition probabilities are $\lambda_1 = 0.2$ and $\lambda_2 = 0.1$, respectively. That is, in expectation, a child spends 4 years as a baby and ten years of school age.

I assume that women's labor force participation can take three values: $h^f \in \{0, 0.2, 0.4\}$, where $h^f = 0$ means no participation, and $h^f = 0.2$ and $h^f = 0.4$

[17]The National Center for Health Statistics estimates that in 2015-19, 14.6% of married women between the ages of 15-29 have impaired fecundity (i.e., who are not surgically sterile, and for whom it is difficult or impossible to get pregnant or carry a pregnancy to term). This number increases to 27.3% and 43.3% when women are in the age range of 30-39 and 40-49, respectively.

[18]In this case, the child is a newborn at age $\bar{J}+1$ and transitions to a teenager 16 years after. If at period $j-1$, the child is not a teenager, the transition to this status occurs with probability one next period.

stand for part-time and full-time work, respectively. The individual's endowment of discretionary time is 5200 hours per year. I consider 16 hours per day, 6 days per week, and 52 weeks per year of discretionary time. A full-time job represents an average of 40 working hours per week, and I assume that part-time workers supply half of the hours. Therefore, full-time workers spend 0.4 of the time endowment in the labor market, while part-time workers spend 0.2 of it.

Wages. To estimate the parameters in the wage equations, I use the Panel Study of Income Dynamics (PSID) data from 1968 to 2015. In particular, I make the same specification for the annual process for log hourly wages as Erosa et al. (2016a):

$$\ln(z_{ij}^g) = \eta^g x_j^g + u_i^g + \nu_j^g + \lambda_j^g, \tag{1.14}$$

where $\ln(z_{ij}^g)$ represents the observed annual log hourly wage of individual i of gender g, at age j in the data, x_j^m and x_j^f represents a second-order polynomial in age and experience, respectively, η^g is the vector of coefficients, $u_i^g \sim N\left(-\frac{\sigma_{u^g}^2}{2}, \sigma_{u^g}^2\right)$ is a fixed effect determined at birth, $\lambda_j^g \sim N\left(-\frac{\sigma_{\lambda,g}^2}{2}, \sigma_{\lambda,g}^2\right)$ is interpreted as measurement error and ν_j^g follows a first-order autoregressive process:

$$\begin{matrix} v_j^m = \rho v_{j-1}^m + \xi_j^m \\ v_j^f = \rho v_{j-1}^f + \xi_j^f \end{matrix} \text{ with } \begin{bmatrix} \xi_j^m \\ \xi_j^f \end{bmatrix} \sim N\left(\begin{matrix} -\frac{\sigma_{\xi^m}^2}{2} \\ -\frac{\sigma_{\xi^f}^2}{2} \end{matrix}, \begin{bmatrix} \sigma_{\xi^m}^2 & 0 \\ 0 & \sigma_{\xi^f}^2 \end{bmatrix} \right), \ v_0^m = v_0^f = 0. \tag{1.15}$$

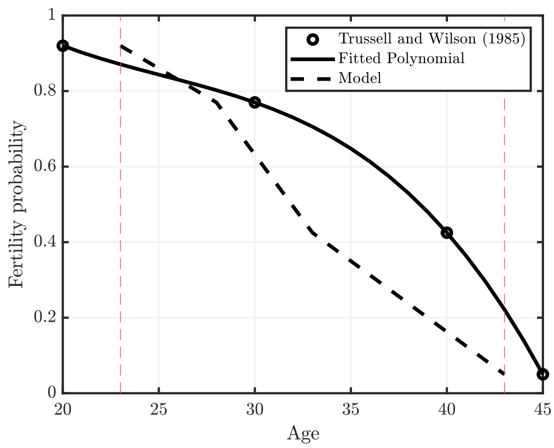
For women, I only take the income shock's standard deviation, the income shock's persistence, and the fixed effect's standard deviation as the parameters estimated in the data. The η^f coefficients in the equation 1.14 are calibrated in the second step to generate some targets regarding the wage gap data. For the numerical solution, I approximate the auto-regressive vector of persistent stochastic shocks for the woman and her partner with a discrete-valued Markov chain using the method proposed by Tauchen (1986) and Tauchen and Hussey (1991).[19] For the fixed effect, instead of discretizing the process, I make a simplifying assumption to reduce the number of state variables. In particular, I assume that the variance of the normal distribution from where the individual draws the realization of the shock is higher for the first period. In the first period, the individual draws a realization of the shock $v_j^g \sim N\left(-\frac{\sigma_{\xi^g}^2}{2}, \sigma_{\xi^g}^2 + \sigma_{u,g}^2\right)$ while from the second period onwards is $v_j^g \sim N\left(-\frac{\sigma_{\xi^g}^2}{2}, \sigma_{\xi^g}^2\right)$. This assumption generates a higher variance in the first periods and, thus, more heterogeneity in the persistent shock. Moreover, while I assume the existence of a transitory shock to estimate the wage process, λ_j^g, I do not model it. In Table 1.4, I list the values of the parameters driving the stochastic process for labor productivity in the baseline economy.

[19] I use five values for the persistent shock.

Tax function. I take the estimations of λ, and τ from Borella et al. (2022). Using the PSID data from 1968 to 2015, they estimate these two by regressing the logarithm of after-tax household income on a constant and on the logarithm of pre-tax household income by year and household type. They provide estimates for single and married households. In this paper, I use their estimates for couples in the year 1982.

Pregnancy probability. I use the same pregnancy probability function as in Erosa et al. (2016b). I take the fertility probability by age when the woman has no child from the estimates in Trussell and Wilson (1985). Specifically, I assume that the fertility probability at age 23 in the model corresponds to the fertility probability at age 20 in the data and that the fertility probability at age 43 in the model corresponds to the fertility probability at age 45 in the data. Figure 1.8 displays the corresponding fertility probabilities in both the model and the data.

FIGURE 1.8 Pregnancy Success Probability Conditional on Age



Note: Author's work, point estimates by Trussell and Wilson (1985). The first and second vertical lines represent the model's first and last fertile ages, respectively.

1.6.2 Second Step

I calibrate the remaining 18 parameters to match 18 data statistics. Table 1.5 reports the values of these calibrated parameters that correspond to the female income process, preference parameters, the pregnancy probabilities, the time cost of children, and the childcare cost. These parameters are selected to target the following data moments that correspond to married women with a college education and belonging to the Middle cohort:

1–3. The proportion of households with zero, one, two, and three children or more.

4. The mother's age at her first child's birth.

5–6. Childcare expenditure as a fraction of average household income for children between 0-4 and 5-15.

7. The average time spent with children

8–10. Employment rates between the ages of 23 and 34 by age groups.

11. The average working hours over the life cycle

12–13. The employment rate and the share of full-time for women who do not have children.

14. The difference in employment rates between mothers who have a child between the ages of 0 and 4 and those who have a child between the ages of 5 and 14.

15. The percentage of full-time workers among mothers who have a child between the ages of 0 and 4

16. The initial gender wage gap between spouses.

17–18. The coefficient of experience and experience squared.

Despite the calibration process implying that all parameters affect all moments, I discuss which parameters provide more information about particular moments.

TABLE 1.5 Calibration: Endogenous Parameters

Parameter	**Value**	**Parameter**	**Value**
Preferences		*Income*	
β	0.910	η_0^f	2.165
γ_c	0.500	η_1^f	0.046
γ_l	0.520	η_2^f	-0.0012
γ_n	0.645	*Pregnancy probability*	
α_c	0.205	$p^{25-29}(0)^*$	0.7
α_l	0.330	$p^{30-34}(0)^*$	0.25
α_n	0.360	$p^{35-39}(0)^*$	0.15
χ	0.190	$p^{40-43}(0)^*$	0.08
ν	0.600	$p^j(1)$	0.88*$p^j(0)$
Children		$p^j(2)$	0.4*$p^j(0)$
ζ_1	0.060		
ζ_2	0.020		
κ_1	0.081		
κ_2	0.058		

Note: * These four parameters are exogenous, however, I show them in this Table for easy reference.

Female earnings. I calibrate η_0^f to match the initial gender wage gap between ages 23-24. The remaining parameters in the earnings equation of women are calibrated by an indirect inference approach. I simulate individual data from the model, I construct the experience accumulation as in the data, and then I run the same Mincer regression as in the data but on model-generated data. I choose η_1^f and η_2^f so that the estimated coefficient of experience and experience squared in the model-generated data equals its analog in the PSID.

Preferences. The utility weights $(\alpha_c, \alpha_l, \alpha_n)$ and the curvatures $(\gamma_c, \gamma_l, \gamma_n)$ are crucial to targeting the female employment rate for childless women, for mothers, and the desire to be a mother for some women, respectively. They are also informative about the share of full-time workers for childless women, the average working hours, and the average life cycle employment profile of women. The direct disutility of being a mother (χ) helps to target the share of women with zero children. The utility of being at home whenever the child is a newborn (ν) is informative about the share of full-time workers of mothers with children between 0-4.

Finally, the discount factor (β) is calibrated to match the age at with women have their first child. This parameter is in the range of the standard values in the literature.[20]

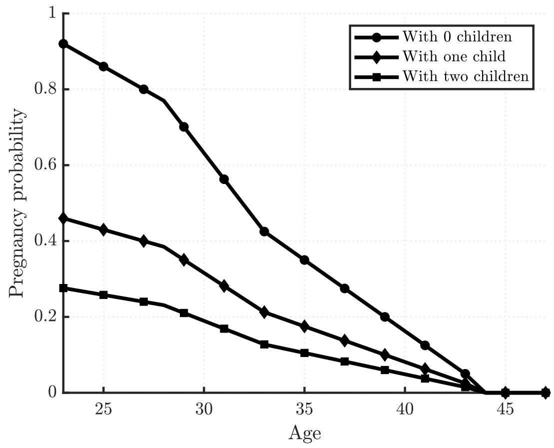
Pregnancy probability. To calibrate the remaining two parameters $(p^j(1)$,$p^j(2))$, I follow the same methodology as in Erosa et al. (2016b). I assume that the fertility probability presented in the previous subsection decreases with the woman's number of children. It measures how difficult it is for a mother to give birth to another child once she has already given birth to one or two children. These parameters are selected to match the share of women with two and three or more children in the model. Furthermore, these pregnancy probabilities are crucial to matching the employment profile of mothers throughout their life. Figure 1.9 shows the pregnancy probabilities as a function of age for women with zero, one, and two children.

Time cost of children. I select the two parameters determining the time cost of children to match the difference in the share of employed mothers of children younger than 5 years old and mothers of children of ages 5-14 and the average time women spend on children. I estimate the time working married women with a college degree spend doing direct childcare using Ramey and Ramey (2009) database. These women in the middle cohort group spend 7.5 hours per week. This implies an average of 390 hours a year.[21]

[20]For example Da Rocha and Fuster (2006) and Attanasio et al. (2008) assume a discount factor of 0.96 and 0.98, respectively.

[21]I add to Ramey and Ramey (2009) database the additional years 2009-19 of the American Time Use Survey. Childcare includes the care of infants, care of older children, medical care of

Figure 1.9 Pregnancy Probability as a Function of Age and Number of Children



Out-of-pocket childcare costs. I assume that the childcare cost for part-time working mothers is half of those who work full-time. Moreover, I calibrate this κ_1 and κ_2 to match the childcare cost implied by children aged between 0-4 and 5-14 in the data. To this end, I use the U.S. Bureau of Census data from the Survey of Income and Program Participation (SIPP) to compute childcare costs.[22] The Census estimates the average childcare payments of working mothers who make childcare payments for the sample period 1993-2011. I compute the average payment for college-educated mothers by the child's age. A child between the ages of 0 and 4 costs about 8 % of the average household income, while it is 5% for children aged between 5 and 14. Hence I calibrate κ_1 and κ_2 to generate a cost of 8% and 5% in the household income, respectively.

1.7 Calibration Results

The Baseline economy is intended to replicate the behavior of college-educated women who marry and belong to the 1944-57 birth cohort. In this section, I present the calibration results of the model. First, I discuss the model fit for the targeted moments. Second, I discuss how informative my model is regarding some untargeted moments. Overall, the model replicates all the targeted moments in the calibration reasonably well.

1.7.1 Targeted Moments

Children-related moments. Figure 1.10 compares the distribution of households regarding the number of children in the model and the data. The model

children, playing with children, helping with homework, reading to and talking with children, dealing with childcare providers, and travel related to childcare.

[22] Guner et al. (2020), and Hannusch et al. (2019) also use these childcare cost estimates.

generates a difference of -0.036, 0.033, 0.007, and -0.005 between the fraction of women with 0, 1, 2, and 3 children in the model and the data, respectively. The age at first birth is 0.7 years later in the model than in the data. These disparities can be attributed to the model's assumption that children are costly in terms of goods. Since households cannot borrow or save, they must delay births more significantly in the model. This effect is particularly pronounced among women with higher initial labor productivity, for whom the opportunity cost of having a child is higher.

Table 1.6 presents how the model behaves regarding the average childcare time mothers spend with their children and the average childcare cost by the child's age and expressed as the share of the average household income. The model captures the monetary cost of children very well, but it underestimates the time women spend a year with their children. This discrepancy could be attributed to the fact that women in the model begin working earlier and are full-time workers compared to the data.

Figure 1.10 Calibration Results: Distribution of Households by Number of Children

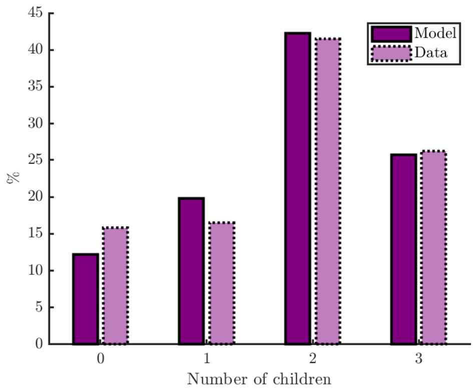


Table 1.6 Calibration Results: Childcare Time and Costs

	Model baseline	**Data baseline**	**Difference**
Average childcare time	355	390	35
Average childcare cost, children 0-4	0.079	0.08	0.001
Average childcare cost, children 5-14	0.046	0.05	0.004

Employment moments. Figure 1.11 compares the female employment rate for seven age groups in the model and the data. I only targeted the employment rate in the first three groups in the calibration. Despite this, the model exhibits a good fit throughout the life cycle. However, the model fails to capture the concave shape observed in the data after age 40. One potential avenue is to incorporate human capital depreciation to achieve a better fit in the latter part of the life cycle. This approach would penalize women primarily due to their absence from work during childbirth, leading to lower incentives to work in the later stages of their life cycle.

Figure 1.11 Calibration Results: Female Employment Rate Over the Life Cycle

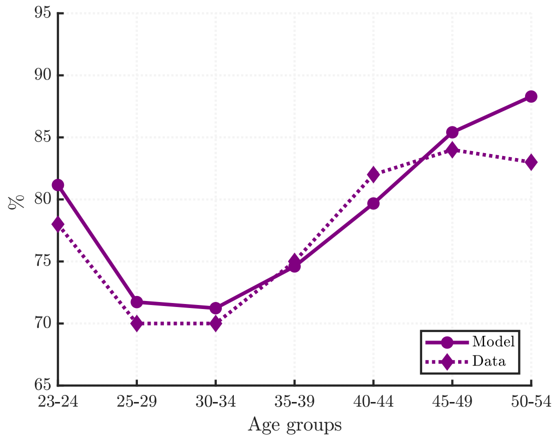


Table 1.7 presents the calibration results for various employment moments. These include the employment rate for childless women, the share of full-time childless women, the share of full-time mothers with children aged 0-4, the employment gap among mothers, and the average working hours. Overall, the model fits the data reasonably well, with most of the differences between the model and data baselines being small. However, childless women's employment and full-time employment rates exhibit larger deviations from the data baseline values, with the differences being -0.069 and -0.079, respectively. Conversely, the model underestimates the average yearly working hours by 107.

Table 1.7 Calibration Results: Employment Moments

	Model Baseline	**Data Baseline**	**Difference**
Employment childless	0.939	0.87	-0.069
Full-time childless	0.749	0.67	-0.079
Employment gap between mothers	0.192	0.22	0.028
Full-time mothers, child 0-4	0.361	0.31	-0.051
Average working hours	1777	1884	107

Note: The employment gap among mothers refers to the difference in the employment rate between mothers whose youngest child is between the ages of 5-14 and those whose youngest child is between 0-4.

Earnings gap and returns to experience. The model replicates well the initial wage gap we observe in the data between ages 24-25. The difference in earnings between men and women in my model is 7.7%, while the difference in data is 6%. Moreover, when regressing log hourly wages of women on a constant, experience, and experience squared in the simulated data from the model, it generates very similar coefficients for the last two when compared with the regression on PSID data, see Table 1.8.

Table 1.8 Calibration Results: Mincer Regression

	Data Baseline	Model Baseline
Experience	0.0282***	0.0294***
	(0.003)	(0.000)
Experience^2	−0.000449***	−0.000403***
	(0.000)	(0.000)
Constant	2.087***	2.231***
	(0.023)	(0.003)
Observations	3894	203950
R-squared	0.0567	0.110

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

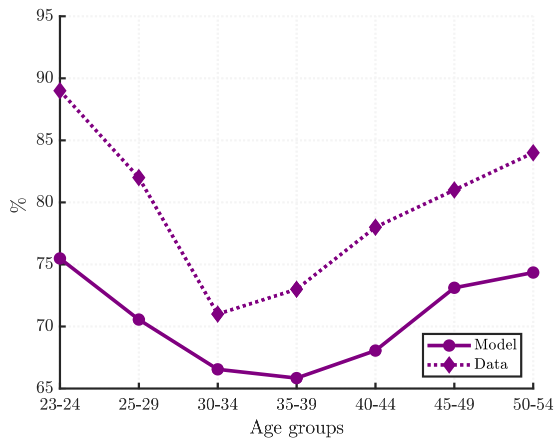
Data baseline: Middle cohort, PSID 1968-2015

Model baseline: simulated data

1.7.2 Untargeted Moments

In the calibration, I target the proportion of full-time employees among non-mothers and the average working hours. However, the share of full-time workers throughout the life cycle is not a targeted moment. Qualitatively, Figure 1.12 illustrates that the model exhibits an appropriate shape over the life cycle. Despite this, the data consistently show a higher proportion of full-time workers across all age groups than the model. Moreover, for mothers, the model produces an employment rate of 73.4%, close to the observed rate of 73% in the data.

Figure 1.12 Untargeted Moments: Share of Women Who Are Full-Time Workers Over the Life Cycle



Furthermore, the model accurately replicates a realistic birth spacing. Specifically, the model indicates a birth spacing of 2.68 years, while the observed spacing in the data is 2.99 years. This shorter spacing in the model is because the average age of women at the time of their first child's birth is higher in the model than in the data. As a result, women who intend to have two or three children must have them in close succession to avoid falling short of their desired number of children due to pregnancy probabilities.

1.8 Quantitative Experiments

This section quantifies the employment and fertility effects resulting from two exogenous changes that impact the decision-making process concerning the choice between pursuing a career and having children. In particular, the two factors are higher marginal returns to experience and the availability of assisted reproductive technology. To achieve this, I perform two experiments. In the first experiment, I calibrate two parameters to increase marginal results on average and at younger ages. In the second experiment, I calibrate the probability of getting pregnant after a certain age to generate the observed increase in the total fertility rate. I refer to these experiments as MRE and MRE+ART, respectively. Finally, I examine how each factor contributes to the sagging middle effect, the postponement of births, and the increase in the total fertility rate observed in the data for the Young cohort.

1.8.1 Increase in Marginal Returns to Experience (MRE)

Measuring the exogenous change across cohorts. I follow two steps to calculate the difference in returns to experience between the Middle and Young cohorts. First, I run a Mincer regression of log hourly wage on experience and experience squared for these two cohort groups. In this regression, the coefficient of experience increases from 0.028 to 0.041; see Table A1.9 in the Appendix. Second, I run the regression for individuals younger than 30 years old. In this case, the coefficient for experience is not statistically significant and equal to 0.00554 for the Middle cohort, while it is statistically significant and equal to 0.0865 for the Young cohort, see Table A1.10 in the Appendix. Thus, returns to experience are greater overall for the Young cohort, and they have an experience premium when they work when young (younger than 30 years old) compared to women in the Middle cohort.

Methodology. In light of these two findings, in this experiment, I first recalibrate the experience coefficient η_1^f to match the overall increase in returns to experience. Second, I change the experience accumulation for the Young cohort. As only for this group of cohorts, individuals younger than 30 ($j \leq 8$) receive an experience premium if they work full-time, I modified the equation 1.7 in the model in the following way:

$$x^f_{j+1} = F(x^f_j, h^f_j) = \begin{cases} x^f_j + 1 + \pi & \text{if } h^f = ft \quad \& \quad j \leq 8 \\ x^f_j + 1 & \text{if } h^f = ft \quad \& \quad j > 8 \\ x^f_j + 0.5 & \text{if } h^f = pt \\ x^f_j & \text{if } h^f = 0. \end{cases} \tag{1.16}$$

The experience premium is captured by π. Since I assume that only women working and younger than 30 years old receive the experiment premium, I make a direct link between female wage growth and birth delays. Women who become mothers at the beginning of their careers and do not work full-time are penalized.

Calibration results. Table 1.9 shows the outcomes of the calibration process and the matched moments. The coefficient related to returns on experience, denoted as η^f_1, exhibits an increase from 0.046 in the baseline model to 0.050 in this experiment. Furthermore, this experiment sets the experience premium, π, to equal 1. Consequently, a woman under 30 years of age, who is employed full-time, accumulates two years of experience, while in the baseline model, she accumulates one year. This model accurately replicates the average marginal returns on experience for women in the Young cohort and the marginal returns to experience for those women below 30 years of age.[23]

Table 1.9 Calibration Results and Targets in MRE Economy

Parameters	Baseline Model	MRE Economy
η^f_1	0.046	0.050
π	0	1
Targets	**Young Cohort**	**MRE Economy**
Marginal returns to experience	0.041	0.039
Marginal returns to experience < 30	0.086	0.077

Note: By "Baseline" I am referring to the model used as a starting point for comparison. "Young cohort" refers to the estimated marginal returns to experience for women born between 1958 and 1978. Lastly, "MRE" stands for the economy where I calibrate the increase in the marginal returns to experience.

Results. In this experiment, I compare two economies: the baseline economy, which replicates the Middle cohort, and the new economy, which aims to reflect the Young cohort's behavior. Table 1.10 shows the average employment and the share of full-time workers for all women, childless women, mothers, and mothers with young children. On average, the increase in returns to experience induces an

[23] I provide in Appendix, Table A1.11, all the coefficients for the two Mincer regressions in the data and the simulated data from this experiment.

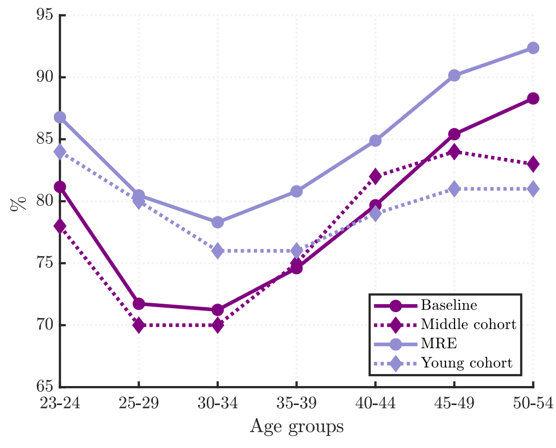
increase in employment of 5.7 pp. Independent of their family status, women find it optimal to work more hours as the opportunity cost of being at home increases. In particular, the highest increase is coming from mothers. Women with children are more likely to participate in the labor market and do so more intensively, as the share of full-time workers also increased by 5.9 pp. Mothers want it all: to be in the labor market and have kids.

TABLE 1.10 Labor Market Outcomes: Baseline vs. MRE

	Baseline	**MRE**	**Difference (pp)**
Employment rate	80.2%	85.9%	5.7
Employment childless	93.9%	95.7%	1.8
Full-time childless	74.9%	80.1%	5.2
Employment mothers	73.4%	77.2%	3.8
Full-time mothers	49.8%	55.6%	5.9
Employment mothers with a child 0-5	60.2%	65.2%	5.0
Full-time mothers with a child 0-5	36.1%	42.6%	6.5

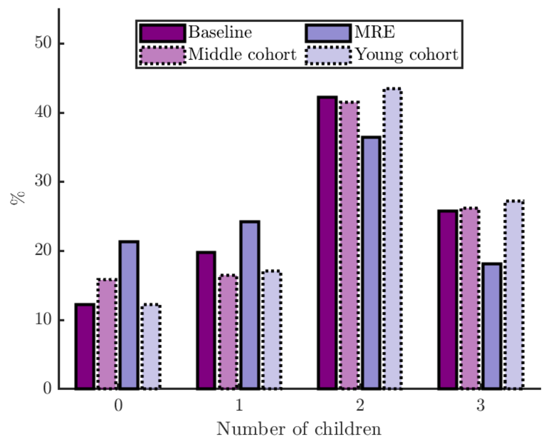
Figure 1.13 compares the employment life-cycle profiles between the baseline and the MRE economies with those of the Middle and Young cohorts. Solid lines illustrate the model outcomes, while the dotted lines represent the data. The results suggest that the increase in returns to experience, particularly at a young age, leads to higher employment rates across all age groups, given the parameters of the baseline economy. The new economy replicates the observed increase in employment rates at ages 23-24. However, it fails to capture the "sagging middle" pattern in the data. Furthermore, the model overestimates the employment rates in this new economy for ages above 30, with the employment rate continuing to rise until it reaches 92%.

FIGURE 1.13 Employment Rate over the Life Cycle in the Baseline and MRE Economies: A Comparison of the Middle and Young Cohorts



Regarding the quantum and timing of fertility, I find that the increase in returns to experience leads to a postponement in the first child's birth by 1.4 years, causing the fertility rate to decrease from 1.8 in the baseline to 1.5. Figure 1.14 compares household distribution by the number of children in the baseline and the MRE economies with those of the Middle and Young cohorts in the data. In the MRE economy, the proportion of women who are not mothers rises from 12.21% to 21.32% relative to the baseline economy. Additionally, the percentage of women with one child increases by 4.36 points, while the shares of women with two and three children decrease by 5.84 and 7.63 percentage points, respectively. The findings demonstrate that an increase in returns to experience, especially at a young age, hurts total fertility rates because women have fewer children and they have them later in life. However, this trend is inconsistent with women's behavior in the Young cohort, suggesting that the trade-off between work and family is more pronounced in this economy than in the data.

Figure 1.14 Distribution of Households by Number of Children in the Baseline and MRE Economies: A Comparison of the Middle and Young Cohorts



On the labor supply side, I find that the increase in returns to experience induces an increase in the employment rate of 5.65 percentage points on average. Moreover, it implies an increase in employment of women between 25 and 29 of 12.2%, which is similar to the increase of 14% in the data. The results of this experiment on the labor supply are qualitatively in line with Olivetti (2006) and Attanasio et al. (2008). They also find that increased returns to experience affect women's labor supply. However, these two papers abstract from modeling fertility decisions and model an increase in returns to experience without the age component that I emphasize in my counterfactual exercise. In this line, my findings are consistent with Caucutt et al. (2002). They use a dynamic general equilibrium model of family formation and investment in children to study the determinants of women's timing of births and labor supply decisions. Although women do not receive a return to labor market experience in the form of higher future wages in their model, they show that women's productivity (wages) delays fertility even when the labor market returns to work experience are zero. My paper complements the previous literature by quantifying how much a change in returns to experience across generations, especially when young, explains the birth delay we observe in the data.

Despite higher returns to experience increasing employment rates at younger ages and delaying births, this exogenous change does not fully explain women's behavior in the Young cohort. This new economy predicts a fall in the total fertility rate by 16.7%, while it actually rises by 4%. I then introduce ART to determine if ART, in conjunction with higher returns to experience, can generate a similar profile of employment and fertility decisions to the Young cohort.

1.8.2 Increase in Marginal Returns to Experience and ART (MRE+ART)

Measuring the increase in ART technology. Given the new income process with higher returns to experience, I additionally change the fertility probabilities to capture the arrival of infertility treatments. It is unclear how much infertility treatments increase the likelihood of having a child. Moreover, most of the ART cycles performed in 2018 were performed on patients under the age of 35.[24] It is still true, however, that infertility increases with age.[25] Therefore, there is a trade-off between ART technology being relevant for older women whose infertility probability is higher and ART technology being used by young women. As a result, I make a conservative assumption: I increase the probability of becoming pregnant for all age groups after age 30 at the same rate $\bar{\pi}$ to match the total fertility rate of the Young cohort.[26]

Table 1.11 summarizes this experiment's calibration results and targets. The model shows that to generate the increase in the total fertility rate of 1.98, the fertility probabilities increase by 0.28 after the age of 30.

Table 1.11 Calibration Results and Targets in the MRE+ART Economy

Parameters	Baseline Model	MRE+ART economy
$\bar{\pi}$	0	0.28
Targets	**Data Young Cohort**	**MRE+ART Economy**
Total fertility rate	1.98	1.97

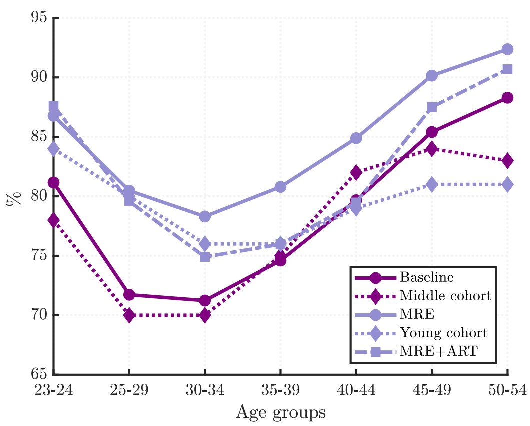
[24]This age group represented 37.2% of all cycles, compared to 22.7% among those aged 35–37, 19.6% among those aged 38–40, 9.4% among those aged 41–42, and 11.1% among those older than age 42.

[25]According to the National Survey of Family Growth, 12.6%, 22.1%, and 26.8% of women aged 15-29, 30-39, and 40-49 were infertile in 2015-2019.

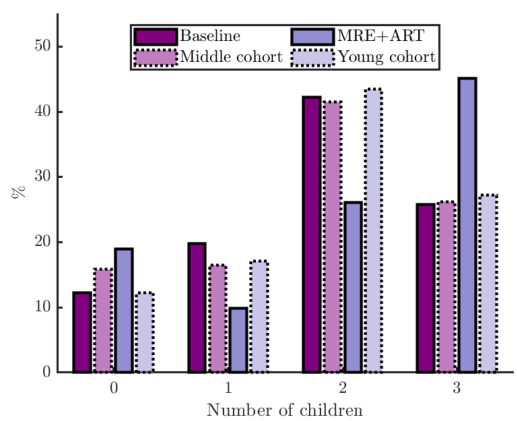
[26]Therefore, I assume that $\tilde{p}^{30-34}(0) = \bar{\pi} + p^{30-34}(0)$, $\tilde{p}^{35-39}(0) = \bar{\pi} + p^{35-39}(0)$ and $\tilde{p}^{40-43}(0) = \bar{\pi} + p^{40-43}(0)$.

Results. In this experiment, I compare the outcomes of the MRE economy and the MRE+ART economy. The former refers to the economy where marginal returns to experience increase compared to the baseline economy, whereas the latter additionally assumes the availability of ART technology. Figure 1.15 presents the results in employment over the life cycle (Panel A) and the distribution of households by the number of children (Panel B). Figure 1.15a shows that when considering ART technology, the employment profile is much more similar to the data of women in the Young cohort. In particular, the employment rate flattens out between ages 30-39 and only increases after this age. Figure 1.15b shows that the relaxation of the opportunity cost of building a career when young, thanks to a greater probability of becoming pregnant when old, increases the share of women with three children by 27 percentage points. As a result, the total fertility rate in the ART economy increases by 0.46 compared to the MRE economy.

Figure 1.15 Employment Rate over the Life Cycle and Distribution of Households by Number of Children in the MRE and MRE+ART Economies Compared to the Middle and Young Cohorts



(a) Employment rate



(b) Distribution of households by children

Using ART technology, new cohorts can relax the trade-off between building their professional career and having a family. Several studies have shown that career and family trade-offs exist. For example, Bertrand et al. (2010) study MBA graduates from the University of Chicago's Booth School of Business. They find a significant trade-off between career and family for US female top professionals. Although this is the case, recent papers like Shang and Weinberg (2013) and Goldin and Katz (2008) indicate that highly educated women in the US have more children and work more.

1.8.3 Discussion

Returns to experience are key for understanding the age profile of employment for the Young cohort. The increased probability of being a mother when old plays a much smaller role in generating this new employment profile. However, fertility outcomes do not match observed data without infertility treatments. Quantitatively, I compare average employment rates, first birth ages, and total fertility rates in the data and those generated by the model to assess the contributions of these two complementary explanations. Table 1.12 summarizes the results.

Table 1.12 Changes in the Data vs. Model

	Data Young vs Middle	**MRE vs Baseline**	**MRE+ART vs Baseline**
Employment rate 25-54	2%	7%	4%
Employment rate 25-29	14%	12%	11%
Employment rate 30-34	9%	10%	5%
Employment rate 35-39	1%	8%	2%
Employment rate 40-44	-4%	7%	0%
Employment rate 45-49	-4%	6%	2%
Employment rate 50-54	-2%	5%	3%
Mother's age first child	8%	5%	6%
Total fertility rate	4%	-17%	9%

The data shows that the employment rate for 25-54-year-old women in the Young cohort is 2% higher than those in the Middle cohort. Despite this minor increase in employment rates, the results of both experiments generate a higher increase than we see in the data. In Experiments MRE and MRE+ART, the employment rate increases by 7% and 4%, respectively. Therefore, both experiments overestimate the change in employment profiles over the life cycle, but the second experiment generates a much similar employment pattern over the life cycle. Regarding the postponement of births, the model generates a much smaller effect than we observe in the data. Compared to the data, the age at which women have their first child increases by 8%, while it is 5% and 6% in Experiment MRE and MRE+ART, respectively. However, only the MRE+ART experiment results in a positive change in total fertility. In the absence of infertility treatments, an increase in returns to experience would result in a 17% decrease in fertility rates, contrary to the 4% increase we observe in the data. As a result of these two exercises, it is evident that returns to experience may play a key role in postponing births. Nonetheless, if infertility treatments had not been available, the fertility rate would have been lower than it is today.

Since this paper aims to quantify how much each exogenous factor, namely an increase in marginal returns to experience and ART, explains employment and fertility behavior among college-educated women born in the Young cohort, the following comparison is made. Table 1.13 shows the percentage difference between the Young cohort data and the experiments. The closer these differences to zero are, the more similar my model's economy is to the Young cohort's. Based on this exercise, the combination of the two factors appears to generate behavior similar to that of the Young cohort.

Table 1.13 Deviation From the Moments of the Young Cohort

	MRE **vs data Young cohort**	**MRE+ART** **vs data Young cohort**
Employment rate 25-54	6%	3%
Employment rate 25-29	1%	-1%
Employment rate 30-34	3%	-1%
Employment rate 35-39	6%	0%
Employment rate 40-44	7%	1%
Employment rate 45-49	11%	8%
Employment rate 50-54	14%	12%
Mother's age first child	0%	1%
Total fertility rate	-24%	0%

1.9 Policy Implications

The new profile of labor supply for college-educated married women seems to have long-lasting effects over the life cycle. While the event of having a kid always discouraged women from work and decreased hours worked, Goldin and Mitchell (2017) show that they take more time out of the labor force after having their kid. Why do these women take more time to return to the labor market compared to previous cohorts? In this paper, I propose a supply-side explanation regarding income taxation. In the United States, the tax system is progressive, and 97% of married couples file their tax return jointly. Filing separate returns as married couples results in a higher tax liability. The main reason relies on the rates applied to taxable income for single filers and married couples.[27] In particular, most tax brackets for married couples are twice the size of those for singles, implying a marriage bonus by filing jointly rather than each partner filing as a single person.

The combination of progressive and joint taxation, where individuals are taxed at the household level, distorts the labor supply decisions of the secondary earner, usually women. This mechanism has been well-studied in the literature.[28] In this

[27]Taxable income is the adjusted gross income minus either the standard deduction or allowable itemized deductions.

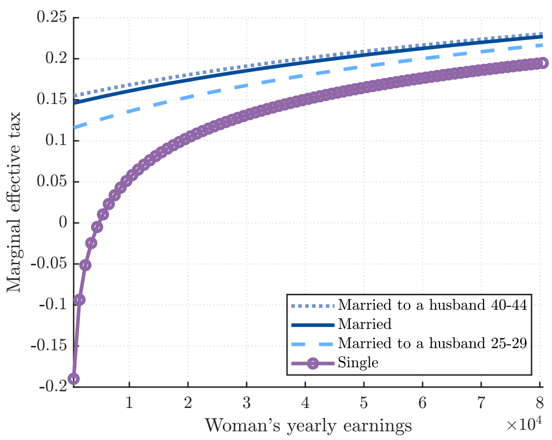
[28]See for example: Guner et al. (2012a), Bronson and Mazzocco (2018), Bick et al. (2019) and Borella et al. (2019).

section, I explore the effect of a shift from a joint to an individual taxation system on the sagging middle effect.

1.9.1 Mechanism

Figure 1.16 illustrates the mechanism I have in mind. It shows the marginal tax rate women pay for different yearly income levels by marital status and the husband's age. Two conclusions can be derived from this simple exercise. First, the marginal tax on single women is lower than that levied on married women. This implies that the change from joint to individual taxation might induce women to participate in the labor market. Second, the female labor supply distortion increases with the husband's age because they tend to have, on average, higher earnings. Therefore the postponement of births combined with joint taxation might explain why women born in the Young cohort group take longer time out of the labor force during the childbearing ages.

Figure 1.16 Female Marginal Tax



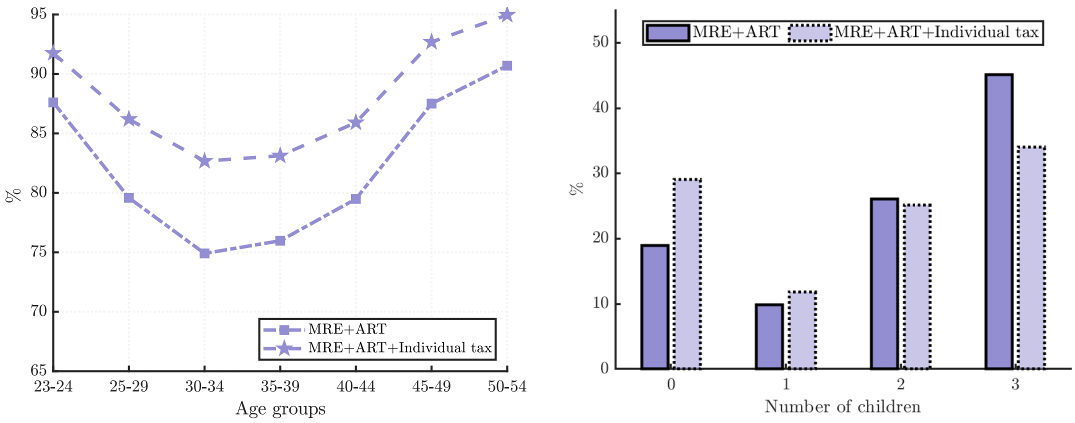
Source: CPS for the mean of the husband's earnings and Borella et al. (2022).

1.9.2 Results of Introducing Individual Taxation

In this section, I run a counterfactual exercise introducing individual taxation. To this end, I take the estimates for the parameters of the tax function for singles from Borella et al. (2022). In particular, $\tau = 0.0787$ and $\lambda = 2.9151$. I also assume that this policy experiment is budget-neutral; therefore, individuals pay a proportional tax (or subsidy) on his/her income.

Figure 1.17 compares the policy reform economy against the economy that accounts for higher marginal returns to experience and assisted reproductive technology. The transition from joint to individual taxation leads to a significant increase in the female employment rate over the life cycle, with an average rise of 3.7%. The increase is particularly high among women aged 30-34, which amounts to 10.4%. This effect is due to fewer children, ultimately decreasing the childcare burden during childbearing. The proportion of childless women rises from 18.9% to 29.02%, while the average age at which women bear their first child increases from 29.14 to 30.22.

FIGURE 1.17 Employment and Fertility Decisions: A Comparison of Joint vs. Individual Taxation



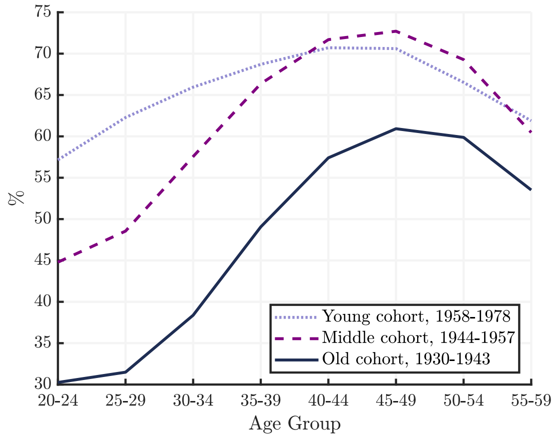
(a) Employment rate (b) Distribution of households by children

1.10 Why Only College-Educated Women?

The new employment life-cycle profile is a particular feature of college-educated married women. Figure 1.18 shows the employment rate of non-college-educated women over their life cycle. At age 20, women in the Young cohort have a higher employment rate than those in the Middle cohort. The employment rates of the Middle and Young cohorts do not become similar until they reach the age of 40. This sample of women does not show a sagging middle effect between the ages of 30-39, as do college-educated women.

At the same time, fertility decisions also differ between college and non-college-educated women. Figure 1.19 shows the age distribution at first birth by cohort and education of the mother. The rise in maternal age for mothers with a college degree is substantial; we can observe how these curves shift to the right, implying an increase in the median age at which these women have their first child. In contrast, women with less education have increasingly delayed childbirth yet have not delayed it to the same extent. The average age at first birth increases by 4.25 and 1.43 for college and non-college women, respectively, between the Old and the Young cohorts. Thus, the delay in fertility has been predominant among women with more

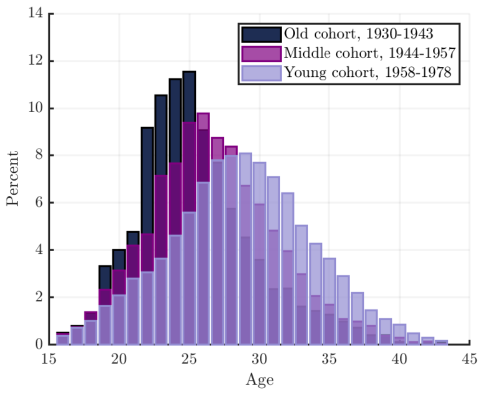
Figure 1.18 Employment Rate of Ever-Married Non-College Educated Women



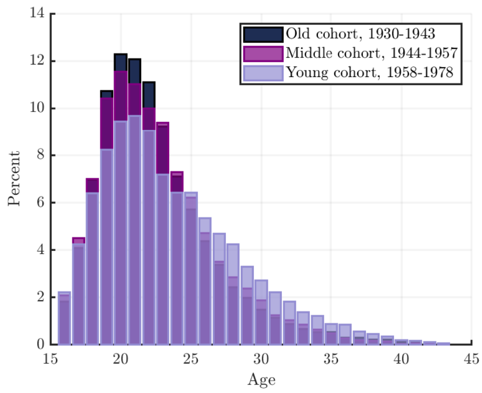
Source: CPS-ASEC microdata, March, 1963 to 2019.
Note: The sample includes ever-married women aged 25-59 without a college education. The variable yearly working hours is computed as the product of last year's weeks and the usual hours worked per week last year.

education.[29] Moreover, while the fertility rate increases by 8pp for college-educated women in the Young cohort compared to the Middle cohort, non-college-educated women keep decreasing the total fertility rate by 9 pp.[30] This section aims to answer the question: why do non-college-educated women behave differently?

Figure 1.19 Age Distribution at First Birth



(a) College



(b) Non-college

Source: CPS June Fertility Supplement (1976-2018).
Note: The sample includes ever-married women with and without a college education. For the mother's age at first birth, I restricted the sample to women older than 16.

[29]See Yang and Morgan (2003) for a deeper analysis of postponement in births by education of the mother.
[30]See Table A1.13 in the Appendix for comparison in the average number of children these two groups of women have as well as the average age at which they have their children.

In this paper, I show the relevant role of higher returns to experience and infertility treatments in generating employment and fertility changes for college-educated women. However, for non-college-educated women, it seems that the mechanism is not that intense. To understand why these two exogenous factors affect non-college-educated women by less, I first analyze the evolution of returns to experience for non-college-educated women. Second, I give several reasons differentiating college and non-college-educated women, which might be relevant to their employment and fertility decisions.

Returns to experience have also increased for non-college-educated women, yet to a lower extent. Table 1.14 shows the results of a Mincer regression where I regress log hourly wages on experience and experience squared. I run this regression for different subsamples: for the Middle, for the Young cohort, and each of those differentiating between college and non-college-educated married women. This analysis shows that returns to experience have increased for all women, independent of their educational attainment. Moreover, in line with the findings for college-educated women, the returns to experience increase when we focus on individuals younger than 30 (see Table A1.12 in the appendix). Despite returns to experience also being higher for those women, this increase is lower than the one for college-educated women. Therefore, the mechanism I highlight in this paper is also present: women have an economic incentive to work more when young and ultimately to postpone fertility. Despite this, the effect for those women is lower.

Table 1.14 Mincer Regression for Married Women by Cohort and Education

	Middle	Middle, NC	Middle, C	Young	Young, NC	Young, C
Experience	0.0333***	0.0306***	0.0311***	0.0643***	0.0536***	0.0625***
	(0.002)	(0.003)	(0.004)	(0.003)	(0.004)	(0.006)
Experience^2	-0.000413***	-0.000351***	-0.000413***	-0.00153***	-0.00104***	-0.00154***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.636***	1.482***	1.974***	1.594***	1.461***	1.875***
	(0.016)	(0.019)	(0.030)	(0.018)	(0.023)	(0.030)
Observations	18977	12504	5278	13633	8379	3629
R-squared	0.0521	0.0558	0.0431	0.0664	0.0704	0.0645

Standard errors in parentheses

Source: PSID 1968-2015

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The increase in the returns to experience for college-educated women could be driven by changes in the occupational structure. Erosa et al. (2017) document that workers in non-linear occupations work on average more hours, and the hourly wage is higher than workers in linear occupations.[31] Over time, the fraction of women working in non-linear occupations (mostly college-educated women) has increased.

[31] Non-linear occupations constitute those occupations with the higher hours group.

Changes in the sorting of college females across occupations might partly explain the increase in the returns to experience relative to non-college-educated women. In Figures A1.8 and A1.9, I provide evidence of a dramatic shift in the distribution of women towards occupations with a higher mean and lower dispersion of hours.

To understand why non-college-educated women postpone births by less than women with a college degree and why they do not have a sagging middle effect, I will propose a set of potential explanations. First, the slope of earnings is flatter for women with lower education.[32] This suggests that women might not find it a sufficient incentive to postpone fertility, even when the return to experience increases. Second, due to the rise in positive assortative mating, (Greenwood et al., 2014), less educated women are less likely to be married to a college-educated husband. Because of this, these women have a lower negative income effect on employment than women married to high-income earners. Finally, for women without a college degree, infertility treatments might not be as relevant as they are for those with one. Two reasons make these treatments less relevant to non-college-educated women's fertility and employment decisions: first, even with the postponement in births, these women have their kids very young and may not require them, and second, these treatments are very expensive and may not be affordable for women without a college degree. Based on Danish administrative data, Groes et al. (2017) finds an education gradient when it comes to IVF success.[33] It is possible that these disparities influence the opportunity cost of delaying childbearing differently among more and less educated women, which may affect fertility choices and labor market outcomes differently.

1.11 Conclusions

A new life cycle of college-educated married women's labor force participation emerged with cohorts born in the mid-1950s. Compared to previous cohorts, their employment profile is flatter and higher with no hump but with a dip in the middle between ages 30-39. In addition, younger cohorts of women are delaying births and increasing their fertility rates.

What brought about the change in women's work and fertility decisions? In this paper, I develop a quantitative theory to provide a unified explanation for the changes across cohorts of employment and fertility decisions of college-educated women. I build a life-cycle model of labor supply and fertility decisions. Children are costly in goods and mother's time. Women face an increased risk of being infertile as they age. These assumptions lead some women to have a low labor supply in

[32]See Figure A1.7 in the Appendix for a comparison in the life-cycle profile of earnings by education.

[33]In the first cycle of pregnancy, Groes et al. (2017) finds that women with a college degree are 24% more likely to deliver a live child than high school dropouts.

their early years until their children grow up and become less expensive. In the model, labor market experience implies higher wages in the future, and the returns to experience are higher for younger women. These assumptions lead some young women to work hard in the market and delay children until later in life. I calibrate the model to match the life-cycle profile of employment, hours worked, and fertility decisions of women born between 1944-1957.

In my model, I assume that two changes in the economic environment trigger the shift in employment and fertility decisions of younger college-educated married women: higher returns to experience, especially at younger ages, and the arrival of Assisted Reproductive Technology (ART). Rising returns to work experience have increased the opportunity cost of career disruptions, especially at the beginning of their careers. Having a child at a young age entails lost wages at this stage and lower human capital accumulation in the long run. If returns to experience are high, women may prefer to postpone childbirth to later ages and invest in a career instead. It means that children will be born during those years when the returns to education materialize, and wages are high. While delaying births increase the risk of infertility, recent advances in ART reduced the risk of infertility, and thus postponing pregnancy has become less costly.

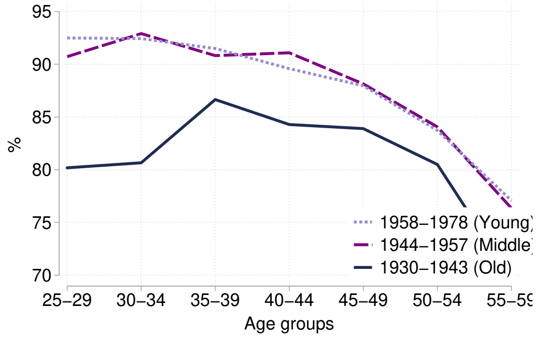
In the first experiment, I take the baseline economy and only increase the returns to experience. Unlike women in the baseline economy, these women have an economic incentive to postpone fertility and work more. As a result, women increase their employment rate between ages 25-29 by 12% and delay the arrival of their first child by 1.38 years. Even though these align with the data, the model predicts a 17% drop in total fertility. In the second experiment, not only do I increase the returns to experience, but I also introduce ART. In this case, women postpone fertility 1.1% more, and their overall employment rate increases by -3.06% less than in the first experiment. According to my findings, the sagging middle effect is primarily caused by higher returns to experience when young. Despite this, ART technology has increased the likelihood of getting pregnant at an older age, which is crucial to capturing the overall increase in fertility we observe. It is, therefore, essential to consider both drivers when analyzing the employment and fertility decisions of young college-educated married women.

The combination of progressive and joint taxation, where individuals are taxed at the household level, distorts the labor supply decisions of the secondary earner, usually women. I show that the change from joint to individual taxation in this context increases the employment rate of women over the life cycle at expenses of a reduction of 16.9% in the total fertility rate.

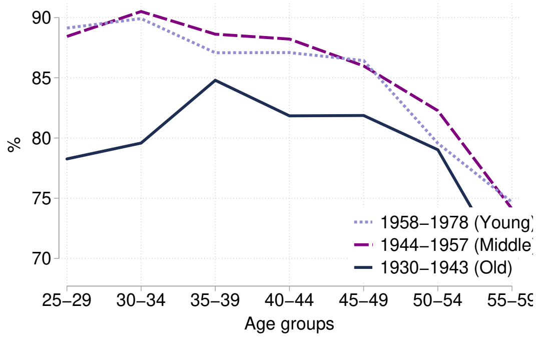
A1 Appendix to Chapter 1

A1.1 Figures

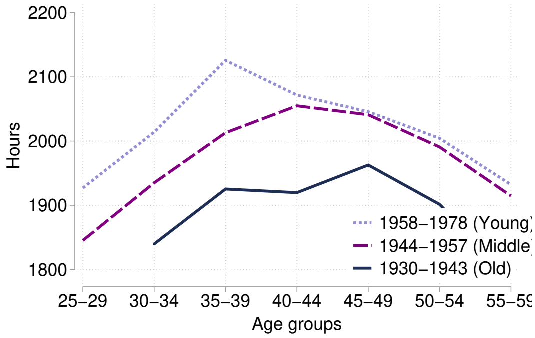
Figure A1.1 Not-Mothers: Married College-Educated Female Labor Market Outcomes



(a) Labor Force Participation



(b) Employment Rate

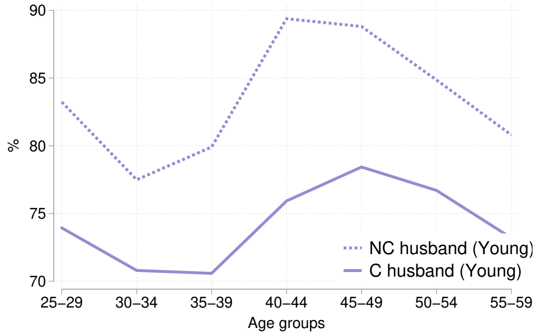


(c) Yearly Working Hours

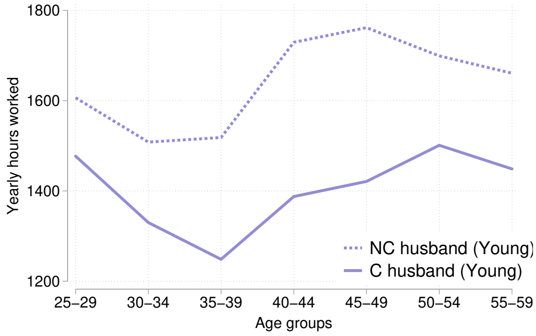
Source: CPS-ASEC microdata, March, 1963 to 2019.

Note: The sample includes married women aged 25-55 with a college education who do not have a child in the household. The variable yearly working hours is computed as the product of the weeks worked last year times the usual hours worked per week last year.

Figure A1.2 Married College-Educated Female Labor Market Outcomes by Husband's Education



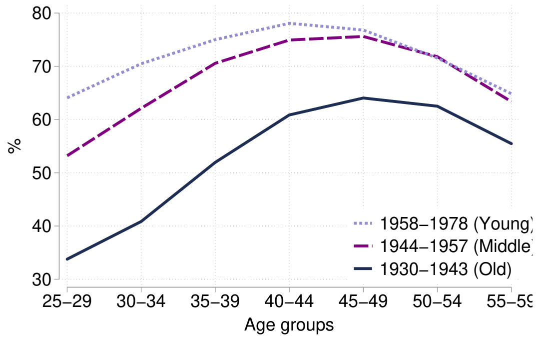
(a) Employment Rate



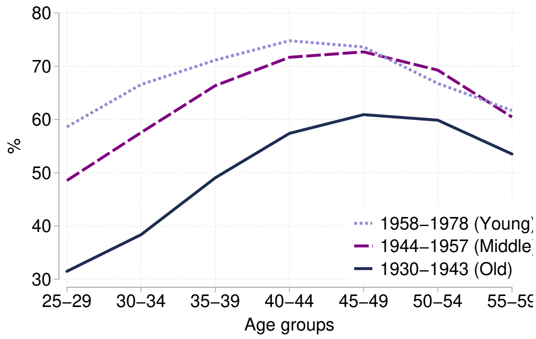
(b) Yearly Working Hours

Source: CPS-ASEC microdata, March, 1963 to 2019.
Note: The sample includes ever-married women aged 25-55 with at least a college education.

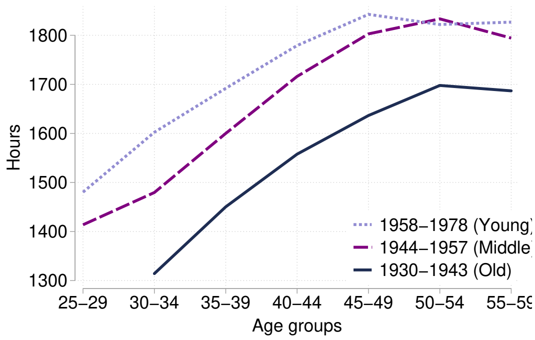
Figure A1.3 Ever-Married Non-College-Educated Female Labor Market Outcomes



(a) Labor Force Participation



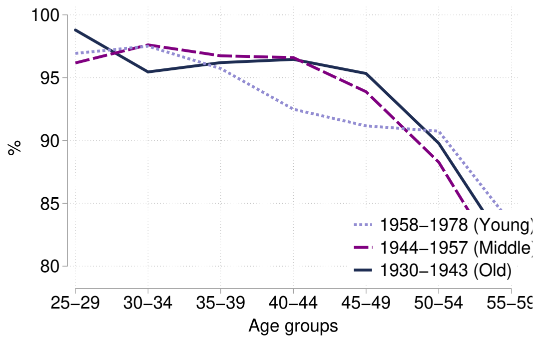
(b) Employment Rate



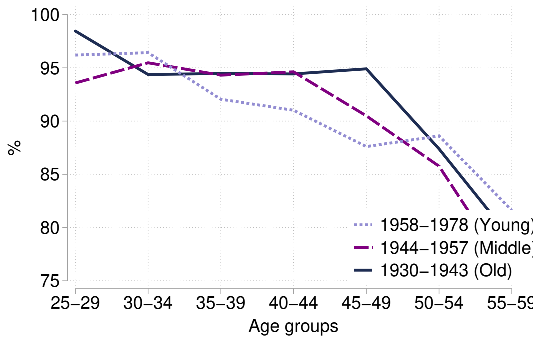
(c) Yearly Working Hours

Source: CPS-ASEC microdata, March, 1963 to 2019.
Note: The sample includes ever-married women aged 25-55 with less than a college education. The variable yearly working hours is computed as the product of the weeks worked last year times the usual hours worked per week last year.

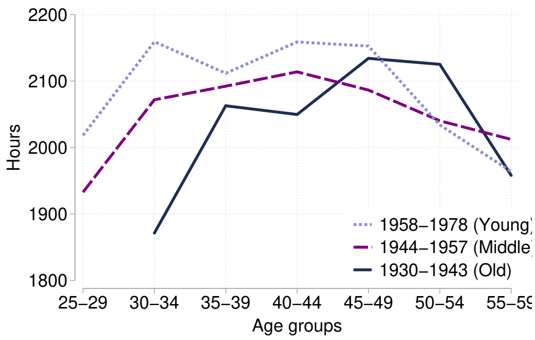
Figure A1.4 Never-Married College-Educated Female Labor Market Outcomes



(a) Labor Force Participation



(b) Employment Rate

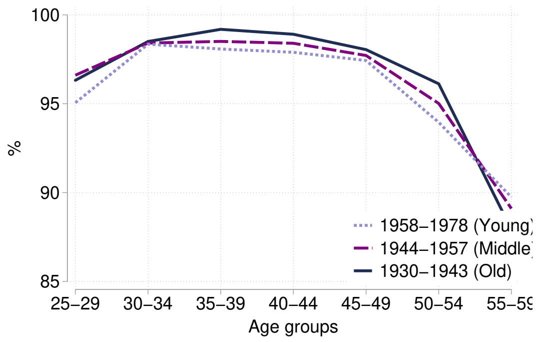


(c) Yearly Working Hours

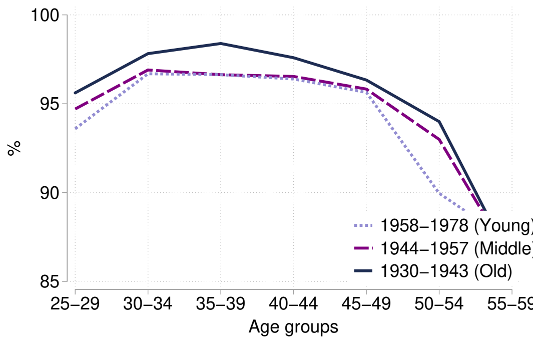
Source: CPS-ASEC microdata, March, 1963 to 2019.

Note: The sample includes never-married women aged 25-55 with a college education. The variable yearly working hours is computed as the product of the weeks worked last year times the usual hours worked per week last year.

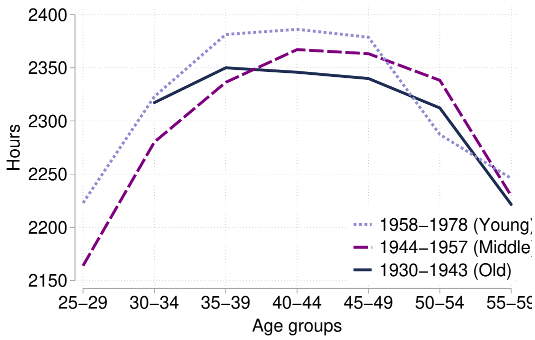
Figure A1.5 Ever-Married College-Educated Male Labor Market Outcomes



(a) Labor Force Participation



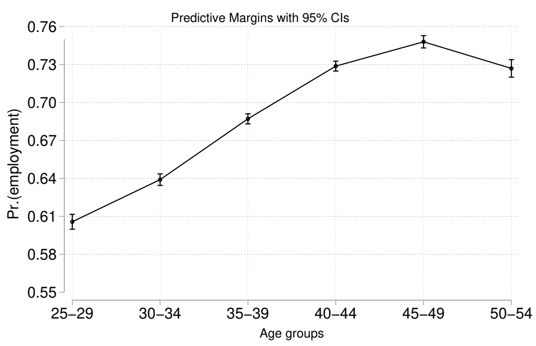
(b) Employment Rate



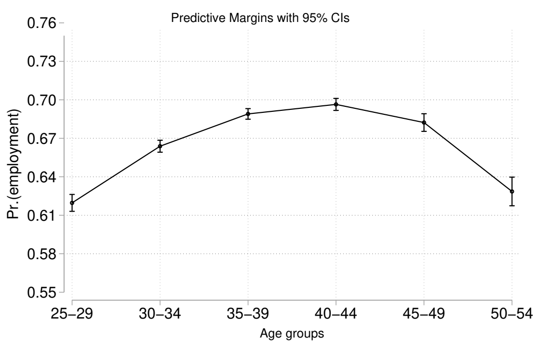
(c) Yearly Working Hours

Source: CPS-ASEC microdata, March, 1963 to 2019.
Note: The sample includes ever-married men aged 25-59 with at least a college education.

Figure A1.6 Marginal Effects for Non-College-Educated Women in Young Cohorts



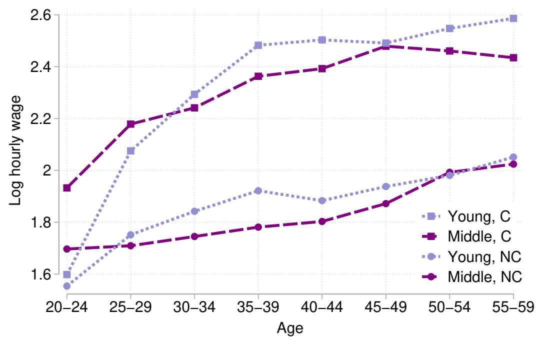
(a) Not Fertility Controls



(b) Fertility Controls

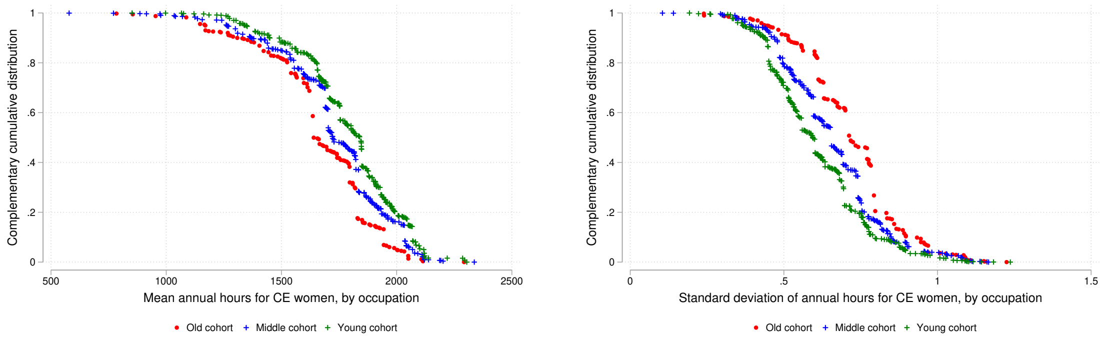
Source: CPS-ASEC microdata, March, 1963 to 2019.
Note: The sample includes ever-married women with less than a college education.

Figure A1.7 Life-Cycle Profile of Earnings by Education



Source: PSID 1968-2015

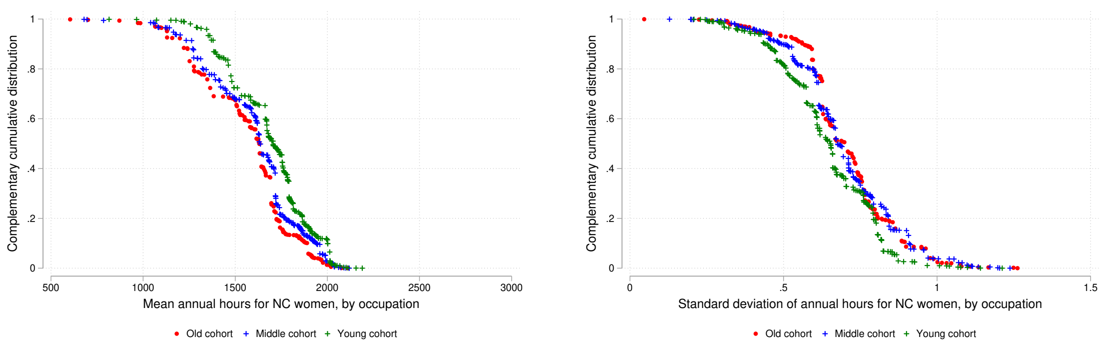
Figure A1.8 Complementary Cumulative Distribution of College-Educated Women by 3-Digit Occupations



(a) Mean Annual Hours Worked (b) Standard Deviation Hours Worked

Source: CPS (1976-2015). Sample: college-educated married women.
Note: The scatter plots show the complementary cumulative distribution of women over occupations in terms of the log of female mean annual hours in an occupation (left panel) or the standard deviation of female log annual hours in an occupation (right panel) in the corresponding time period.

Figure A1.9 Complementary Cumulative Distribution, Non-College-Educated Women: by 3-Digit Occupations



(a) Mean Annual Hours Worked

(b) Standard Deviation Hours Worked

Source: CPS (1976-2015). Sample: non-college-educated married women.

Note: The scatter plots show the complementary cumulative distribution of women over occupations in terms of the log of female mean annual hours in an occupation (left panel) or the standard deviation of female log annual hours in an occupation (right panel) in the corresponding time period.

A1.2 Tables

Table A1.1 Estimates of the Probability of Employment for the Young Cohort

	Model (I)	Model (II)	Model (III)	Model (IV)
	dy/dx	dy/dx	dy/dx	dy/dx
Age groups=30-34	-0.0356***	-0.00238	0.0115*	0.0246***
	(-7.62)	(-0.49)	(2.25)	(4.12)
Age groups=35-39	-0.0289***	0.0144**	0.0429***	0.0370***
	(-5.93)	(2.78)	(7.97)	(5.86)
Age groups=40-44	0.00575	0.0465***	0.0782***	0.0370***
	(1.15)	(8.82)	(14.21)	(5.36)
Age groups=45-49	0.0247***	0.0537***	0.0831***	0.00924
	(4.63)	(9.54)	(14.26)	(1.15)
Age groups=50-54	0.0235***	0.0210**	0.0495***	-0.0384***
	(3.84)	(3.16)	(7.29)	(-3.81)
College educated wife=1	-0.0653***	-0.0641***	-0.0632***	-0.0659***
	(-26.17)	(-25.78)	(-25.51)	(-22.76)
Husband's income	-0.0878***	-0.0836***	-0.0809***	-0.0917***
	(-47.06)	(-45.20)	(-44.13)	(-43.30)
Husband's full-time-full-year indicator=1	0.0188***	0.0216***	0.0225***	0.0120*
	(4.35)	(4.99)	(5.21)	(2.42)
Mother=1		-0.115***	-0.0451***	0.0966***
		(-47.41)	(-13.07)	(10.62)
Number of children			-0.0481***	-0.0466***
			(-39.60)	(-34.53)
Age youngest child				0.0127***
				(33.70)
Observations	142702	142702	142702	115538
Year	Yes	Yes	Yes	Yes
Husband's work type	Yes	Yes	Yes	Yes
Native indicator	Yes	Yes	Yes	Yes

t statistics in parentheses

Source: March CPS

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This Table computes the marginal probability for the college-educated women belonging to the Young cohorts.

Table A1.2 Estimates of the Probability of Employment for All Cohorts - Model (I)

	Old	Middle	Young
	dydx	dydx	dydx
Age groups=30-34	-0.0462***	-0.0452***	-0.0356***
	(-3.37)	(-7.14)	(-7.62)
Age groups=35-39	0.00773	-0.0248**	-0.0289***
	(0.49)	(-3.26)	(-5.93)
Age groups=40-44	0.0593**	0.0323***	0.00575
	(3.27)	(3.67)	(1.15)
Age groups=45-49	0.0726***	0.0502***	0.0247***
	(3.57)	(5.03)	(4.63)
Age groups=50-54	0.0487*	0.0465***	0.0235***
	(2.09)	(4.14)	(3.84)
College educated husband=1	-0.101***	-0.0593***	-0.0653***
	(-12.66)	(-16.23)	(-26.17)
Husband's income	-0.102***	-0.0924***	-0.0878***
	(-10.93)	(-28.05)	(-47.06)
Husband's full-time-full-year indicator=1	0.0730***	0.0343***	0.0188***
	(4.86)	(5.93)	(4.35)
Observations	21347	70674	142702
Year	YES	YES	YES
Husband's work type	YES	YES	YES
Native indicator	YES	YES	YES

t statistics in parentheses

Source: March CPS

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This Table computes the marginal probability for the college-educated women belonging to the Old, Middle and Young cohorts.

Table A1.3 Estimates of the Probability of Employment for All Cohorts - Model (IV)

	Old	Middle	Young
	dydx	dydx	dydx
Age groups=30-34	0.0301	0.0215**	0.0246***
	(1.73)	(2.93)	(4.12)
Age groups=35-39	0.0442*	0.0307***	0.0370***
	(2.32)	(3.43)	(5.86)
Age groups=40-44	0.0280	0.0307**	0.0370***
	(1.26)	(2.77)	(5.36)
Age groups=45-49	-0.0116	-0.00964	0.00924
	(-0.47)	(-0.70)	(1.15)
Age groups=50-54	-0.101***	-0.0617***	-0.0384***
	(-3.57)	(-3.77)	(-3.81)
Mother=1	0.118***	0.0798***	0.0966***
	(7.97)	(8.05)	(10.62)
Number of children	-0.0174***	-0.0361***	-0.0466***
	(-5.15)	(-17.72)	(-34.53)
Age youngest child	0.0242***	0.0168***	0.0127***
	(24.13)	(30.97)	(33.70)
College educated husband=1	-0.0845***	-0.0633***	-0.0659***
	(-9.33)	(-14.42)	(-22.76)
Husband's income	-0.134***	-0.0991***	-0.0917***
	(-14.26)	(-25.34)	(-43.30)
Husband's full-time-full-year indicator=1	0.0783***	0.0347***	0.0120*
	(4.72)	(5.07)	(2.42)
Observations	16026	52596	115538
Year	YES	YES	YES
Husband's work type	YES	YES	YES
Native indicator	YES	YES	YES

t statistics in parentheses

Source: March CPS

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.4 Fertility Decisions of College-Educated Married Women by Cohort

Cohort	Number of children	Age 1st child	Not mothers at age 40
1930-43, Old	2.51	25.4	10.0%
1944-57, Middle	1.90	26.7	16.0%
1958-78, Young	1.98	28.9	12.0%

Source: CPS June Fertility Supplement (1976-2018)

Table A1.5 Fertility Decisions of College-Educated Married Women by Cohort and Husband's Education

	Number of children		Age 1st child	
Cohort	NC husband	C husband	NC husband	C husband
1930-43, Old	2.46	2.46	24.8	25.7
1944-57, Middle	1.89	1.91	25.8	27.2
1958-78, Young	1.88	2.01	27.3	29.4

Source: CPS June Fertility Supplement (1976-2018)

Table A1.6 Mincer Regression for College-Educated Married Women by Cohort (IV)

	Middle cohort	Young cohort
Age dummy (< 30)	-0.227***	-0.329***
	(0.046)	(0.059)
Experience	0.0150***	0.00967**
	(0.003)	(0.003)
Age dummy (< 30) × Experience	0.000325	0.0283***
	(0.005)	(0.008)
Number of children	-0.0595***	0.0172
	(0.015)	(0.014)
Husband's education	0.00912	0.0569***
	(0.009)	(0.007)
Husband's hourly wage	0.152***	0.112***
	(0.032)	(0.024)
Husband's FT contract dummy	0.0704*	0.131***
	(0.029)	(0.036)
FT contract dummy	0.0140	-0.233**
	(0.041)	(0.074)
Husband's experience	-0.00767***	0.00301
	(0.002)	(0.003)
Constant	1.734***	1.152***
	(0.157)	(0.147)
Observations	4926	3376
R-squared	0.0859	0.126

Standard errors in parentheses

Source: PSID 1968-2015

Sample: College-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.7 Mincer Regression for College-Educated Married Women by Cohort (V)

Age dummy (< 30)	-0.162***
	(0.042)
Experience	0.0140***
	(0.002)
Age dummy (< 30) × Experience	-0.00177
	(0.004)
Dummy Young cohort	0.139**
	(0.049)
Dummy Young cohort × Age dummy (< 30)	-0.303***
	(0.068)
Dummy Young cohort × Experience	-0.00347
	(0.003)
Dummy young cohort × Age dummy (< 30) × Experience	0.0378***
	(0.010)
Constant	2.121***
	(0.032)
Observations	9124
R-squared	0.0536

Standard errors in parentheses

Source: PSID 1968-2015

Sample: College-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.8 Mincer Regression for College-Educated Married Women by Cohort (VI)

	Middle cohort	Young cohort
Age dummy (< 30)	-0.162***	-0.466***
	(0.042)	(0.053)
Experience	0.0140***	0.0105***
	(0.002)	(0.003)
Age dummy (< 30) × Experience	-0.00177	0.0360***
	(0.004)	(0.009)
Constant	2.121***	2.261***
	(0.032)	(0.037)
Observations	5420	3704
R-squared	0.0412	0.0725

Standard errors in parentheses

Source: PSID 1968-2015

Sample: College-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.9 Mincer Regression for College-Educated Married Women by Cohort (VII)

	Middle	Young
Experience	0.0282***	0.0413***
	(0.003)	(0.006)
Experience^2	-0.000449***	-0.000949***
	(0.000)	(0.000)
Constant	2.087***	2.082***
	(0.023)	(0.035)
Observations	3894	2720
R-squared	0.0567	0.0447

Standard errors in parentheses

Source: PSID 1968-2015

Sample: college-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.10 Mincer Regression for College-Educated Married Women by Cohort (VIII)

	<30	>30
Experience	0.00554	0.0309***
	(0.009)	(0.004)
Experience^2	0.000183	-0.000554***
	(0.000)	(0.000)
Constant	2.135***	2.091***
	(0.033)	(0.032)
Observations	740	2972
R-squared	0.00881	0.0461

Standard errors in parentheses
Source: PSID 1968-2015
Sample: college-educated married women
Note: < 30 includes women below 30
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	<30	>30
Experience	0.0865***	0.0218**
	(0.026)	(0.008)
Experience^2	-0.00793***	-0.000488
	(0.002)	(0.000)
Constant	1.926***	2.267***
	(0.066)	(0.056)
Observations	822	2022
R-squared	0.0168	0.0104

Standard errors in parentheses
Source: PSID 1968-2015
Sample: college-educated married women
Note: < 30 includes women below 30
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The left side corresponds to the Middle cohort, while the right side corresponds to the Young cohort.

Table A1.11 Mincer Regression for College-Educated Married Women by Cohort (IX)

	Young cohort	Experiment 1
Experience	0.0413***	0.0389 ***
	(0.006)	(0.000)
Experience^2	-0.000949***	-0.000677***
	(0.000)	(0.000)
Constant	2.082***	2.2685***
	(0.035)	
Observations	2720	236260
R-squared	0.0567	0.148

Standard errors in parentheses
Source: PSID 1968-2015
Sample: college-educated married women

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	Young cohort	Experiment 1
Experience	0.0865 ***	0.0771 ***
	(0.026)	(0.006)
Experience^2	-0.00793***	0.0038
	(0.002)	(0.001)
Constant	1.926***	2.15***
	(0.066)	(0.008)
Observations	822	33998
R-squared	0.0168	0.149

Standard errors in parentheses
Source: PSID 1968-2015
Sample: college-educated married women
younger than 30 years old
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The left side corresponds to All Sample Working Women, while the right side corresponds to Below 30 Years Old.

Table A1.12 Mincer Regression for Married Women by Cohort and Education (II)

	Middle	Middle, NC	Middle, C	Young	Young, NC	Young, C
Age dummy (< 30)	-0.00868	0.0170	-0.0843	-0.482***	-0.372***	-0.529***
	(0.033)	(0.039)	(0.059)	(0.047)	(0.061)	(0.070)
Experience	0.0339***	0.0328***	0.0279***	0.0282***	0.0222***	0.0281***
	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)	(0.008)
Age dummy (< 30) × Experience	-0.00997	0.0103	-0.0124	0.118***	0.0940***	0.135***
	(0.006)	(0.007)	(0.010)	(0.014)	(0.016)	(0.024)
Experience^2	-0.000456***	-0.000388***	-0.000391**	-0.000560***	-0.000174	-0.000644*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age dummy (< 30) × Experience^2	-0.000236	-0.00110**	0.000190	-0.00772***	-0.00467***	-0.0130***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Constant	1.662***	1.464***	2.035***	1.883***	1.692***	2.171***
	(0.026)	(0.031)	(0.048)	(0.038)	(0.049)	(0.057)
Observations	19491	12850	5420	13924	8561	3704
R-squared	0.0516	0.0539	0.0433	0.0682	0.0656	0.0800

Standard errors in parentheses
Source: PSID 1968-2015
Sample: Married women
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A1.13 Fertility Decisions of College and Non-College Educated Women by Cohort

	College	**Non-college**	**College**	**Non-college**
Cohort	**Number of children**		**Age at first birth**	
Old	2.51	3.12	24.75	22.44
Middle	1.90	2.32	26.42	22.58
Young	1.98	2.23	29.00	23.87

Source: CPS June Fertility Supplement (1976-2018)

A1.3 Computational Appendix

Stochastic structure for children's ageing

In this Appendix, I show how to retrieve the stochastic structure that governs the transition probabilities for the vector state of the number of children at different ages takes, which I denote by $\mathbf{n}' = \Lambda_j(\mathbf{n}, b)$.

I denote by λ_1 and λ_2 the probabilities that an individual baby becomes a school-age child, and that an individual school-age child becomes a teen in a given period, respectively. Moreover, I assume that the aging event is independent across children. Notice that $bp^j(N)$ is the probability that there is a newborn in the next period. Denote by $P_i(x \mid n_i)$ the probability that x children in stage $i \in \{1, 2\}$ (babies, school-age) move on to the next stage the next period (school-age, teenager), conditional on there being n_i children in that stage in the current period. Table A1.14 shows these, for babies and school-age children.

Table A1.14 Probabilities of Aging by Number of Children, $P_i(x \mid n_i)$

n_i	Number of children ageing			
	0	1	2	3
0	1	0	0	0
1	$1-\lambda_i$	λ_i	0	0
2	$(1-\lambda_i)^2$	$\lambda_i(1-\lambda_i)$	λ_i^2	0
3	λ_i^3	$\lambda_i(1-\lambda_i)^2$	$\lambda_i^2(1-\lambda_i)$	λ_i^3

To compute the whole set of probabilities of transition from one state to the other, $\mathbf{n} = [n_0, n_1, n_2, n_3]$, I follow the algorithm, where $\bar{P} = P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$:

for $x_1 \in \{0, 1, 2, 3\}$ **do**
  for $x_2 \in \{0, 1, 2, 3\}$ **do**
    if $n_0 = 1$ **then**
      $\mathbf{n}' = [0, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $(1 - bp^j(N))\bar{P}$ or
      $\mathbf{n}' = [1, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $bp^j(N)\bar{P}$;
    else
      $\mathbf{n}' = [0, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $(1 - bp^j(N))\bar{P}$ or
      $\mathbf{n}' = [1, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $bp^j(N)\bar{P}$;
    end
  end
end

Taste Shocks

All of the decisions women make in the model are discrete. To facilitate the numerical solution of the model, I include a taste shock to women's utility in every period. This helps by smoothing out labor force participation and fertility decisions. The shocks can be interpreted as unobserved state variables that add noise to the women's decisions. Moreover, the calibration and results are robust to their inclusion. For an in-depth discussion of this computational method, see Iskhakov et al. (2017).

Thus, I assume that in every period, women receive a vector of additive-separable taste shocks μ. In periods when they can still have children and need to choose on pregnancy $b \in \{0, 1\}$ in addition to labor force participation $h^f \in \{0, \frac{1}{4}, \frac{1}{2}\}$, they receive a vector of six shocks, one for every element in $\{0, \frac{1}{4}, \frac{1}{2}\} \times \{0, 1\}$. In periods when they cannot have any more children and need only to choose labor force participation, they receive a vector of three shocks, one for every element in $\{0, \frac{1}{4}, \frac{1}{2}\}$:

$$\mu = \begin{cases} \left(\mu_{0,0}, \mu_{\frac{1}{4},0}, \mu_{\frac{1}{2},0}, \mu_{0,1}, \mu_{\frac{1}{4},1}, \mu_{\frac{1}{2},1}\right) & \text{if } j < \bar{J} \text{ and } N < 3 \\ \left(\mu_0, \mu_{\frac{1}{4}}, \mu_{\frac{1}{2}}\right) & \text{otherwise} \end{cases}$$

All of these shocks are i.i.d, drawn from an Extreme Value Type I distribution with scale parameter σ_μ. The modified value function in states $\left(x^f, \mathbf{n}, v^m, v^f\right)$ is:

$$W_j\left(x^f, \mathbf{n}, v^m, v^f, \mu\right) = \begin{cases} \max\{W_j^{h^f,b}\left(x^f, \mathbf{n}, v^m, v^f\right) + \sigma_\mu \mu_{h^f,b}\}_{h^f \in \{0,\frac{1}{4},\frac{1}{2}\}, b \in \{0,1\}} & \text{if } j < \bar{J} \text{ and } N < 3 \\ \max\{W_j^h\left(x^f, \mathbf{n}, v^m, v^f\right) + \mu_h\}_{h^f \in \{0,\frac{1}{4},\frac{1}{2}\}} & \text{otherwise,} \end{cases}$$

where $W_j^{h^f,b}$ and $W_j^{h^f}$ represent the value, ex-taste shock, of choosing labor force participation h^f and pregnancy status b for a woman in period j, or just labor force participation h^f, in states $\left(x^f, \mathbf{n}, v^m, v^f\right)$:

$$\begin{aligned} W_j^{h^f,b}\left(x^f, \mathbf{n}, v^m, v^f\right) &= u^{h^f,b}(c, l, \mathbf{n}) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(x'^f, \mathbf{n}', v'^m, v'^f\right)\right] \\ W_j\left(x^f, \mathbf{n}, v^m, v^f\right) &= u^{h^f}(c, l, \mathbf{n}) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(x'^f, \mathbf{n}', v'^m, v'^f\right)\right], \end{aligned}$$

where $\mathbb{E}^{\sigma_\mu}$ denotes the expectations over future taste shocks, and in both cases, choice and state variables need to be retrieved from the constraints and laws of motion.

The main consequence of introducing the taste shocks is that the policy function becomes probabilistic. Given the distribution assumed for them, the probability that a woman chooses pregnancy decision b and labor force participation h^f in states $\left(x^f, \mathbf{n}, v^m, v^f\right)$ when $j < \bar{J}$ and $N < 3$ is the logit probability:

$$P_j\left(h^f, b \mid x^f, \mathbf{n}, v^m, v^f\right) = \frac{\exp\left(\frac{W_j^{h^f,b}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}}\sum_{k\in\{0,1\}}\exp\left(\frac{W_j^{i,k}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)}.$$

Otherwise, the probability that a woman chooses labor force participation h in states $\left(x^f, \mathbf{n}, v^m, v^f\right)$ is the logit probability:

$$P_j\left(h^f \mid x^f, \mathbf{n}, v^m, v^f\right) = \frac{\exp\left(\frac{W_j^{h^f}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}}\exp\left(\frac{W_j^{i}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)}.$$

One additional benefit of using Extreme Value Type I shocks is that the expected value function is given by the tractable log-sum formula from (McFadden, 1973):

$$\mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(x'^f, \mathbf{n}', v'^m, v'^f\right)\right] =$$
$$\begin{cases}\sigma_\mu \log\left(\sum_{i\in 0,1,2,3}\sum_{k\in 0,1}\exp\left(\frac{W_j^{i,k}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)\right) & \text{if } j < \bar{J} \text{ and } N < 3\\ \sigma_\mu \log\left(\sum_{i\in 0,1,2,3}\exp\left(\frac{W_j^{i}\left(x^f,\mathbf{n},v^m,v^f\right)}{\sigma_\mu}\right)\right) & \text{otherwise.}\end{cases}$$

Using backward induction starting in period J, one can easily retrieve the expected value functions and the probabilistic policy functions.

Chapter 2

Cash Transfers and Fertility: From Short to Long Run

2.1 Introduction

In the developed world, the prevailing total fertility rates (TFR) have been well under the replacement level of 2.1 since the mid-1980s. This implies, in the absence of immigration, that the total population has already or will soon start to decline and that the share of old people will rise. These trends are expected to have harmful effects on economic performance.[1] Governments worldwide, worried about demographic pressures and their effects on the economy, are deploying policies to boost fertility rates and slow down population decline and aging. There may be room for policy intervention due to the large gap between desired and realized birth rates among recent cohorts of women. Can family policies effectively close this gap and thus increase the number of children born per woman?

To answer this question, we propose a dynamic life-cycle model featuring joint fertility and labor force participation decisions. To calibrate the model, we use a natural experiment involving a policy that gave one-shot cash transfers to women upon childbirth in Spain between 2007-2010, known in the media as the "cheque bebé" (baby check). We choose the model's parameters to match its effects on fertility and labor force participation in the year following implementation (the short

[1]Population aging and decline are associated in the literature with increasing burdens of social security systems (De Nardi et al., 1999), low-interest rates (Krueger and Ludwig, 2007) and reduced output growth and investment (Aksoy et al., 2019). Jones (2022) explores the consequences of population decline in growth models. A disturbing possibility emerges: the Empty Planet result. This refers to a situation in which knowledge and living standards stagnate for a vanishing population, which can happen if society doesn't implement the optimal allocation fast enough. He concludes that policies related to fertility can determine whether we fall into this trap or converge to a balanced growth path with exponential growth ("Expanding Cosmos").

run).[2] These responses contain invaluable information regarding the sensitivity of such decisions to family income. But they may also reflect changes in fertility timing (*tempo*) induced by the policy that does not necessarily translate into a change in the completed fertility rate. We also use data on desired fertility to capture heterogeneity in preferences over the number of children and match overall labor supply patterns for groups of women with children of different ages, as well as the average timing of first births right before the policy implementation.

We then use the calibrated model as a laboratory to perform quantitative experiments. First, we simulate the effects of the policy in the long run, i.e., after enough time has passed so that all women have been exposed to cash transfers since the beginning of their childbearing years. Thus, we quantify the effects on the overall number of children per woman (*quantum*). Then, we simulate the effects of an alternative policy involving childcare subsidies for children aged 0-3, with the same present value as the cash transfers. The cost of childcare at these ages can play an important role in fertility and mothers' labor force participation decisions since public schooling is usually not universally available yet for these children.[3] Finally, we explore the stand-alone effects and the interactions with the cash transfers arising from an essential feature of Spanish labor markets: the coexistence of temporary and permanent contracts, known as the duality of the labor market. Most workers start their careers with temporary contracts, which generate lower yearly earnings (partly due to intermittent unemployment) and have lower returns to experience. Therefore, the costs of career interruptions may be substantial in crucial child-bearing years, as women want to work, accumulate experience and increase the likelihood of transitioning to a permanent contract.

Our main result is that the long-run effects of cash transfers on fertility are about half as large in magnitude as the short-run effects (3% and 6%, respectively), the main reason being that in the short run, there are additional births among older cohorts of women that only have a few periods to adjust their decisions after the policy is announced. Moreover, we find that childcare subsidies for parents of children aged 0-3, with the same present value as the cash transfers considered in the first exercise, have a long-run impact on fertility that is only slightly smaller than the latter. However, instead of reducing mothers' labor force participation, which the cash transfers do, they increase it (around one percentage point, while the unconditional transfers reduce it in a magnitude at least as large). Furthermore, we find that eliminating the dual nature of Spanish labor markets (i.e., implementing a unique contract with earnings and returns to experience between the temporary and the permanent ones) has a long-run positive effect on fertility which is three times as large as that of the cash transfers (9% increase). Finally, we explore the

[2]See González (2013).

[3]This was indeed the case in Spain at the time when cash transfers were introduced, see section 2.2 for details.

interaction between labor market duality and cash transfers. In the long run, we find that the latter's effects on fertility are very similar both in a scenario with dual labor markets and in one where there is a single contract (which is more representative of the situation in other developed countries like the United States).

Taken together, our results imply that the impact of family policies may be smaller in the long than in the short run since the mechanisms at work do not seem specific to the cash transfer policy we consider here. Moreover, these quantitative exercises highlight the importance of interventions that make labor market interruptions less costly (as eliminating duality) and less necessary (as childcare subsidization) for fertility outcomes when considering the decision jointly with labor force participation.

Related literature and contributions.— This paper is related to various lines of research that intersect with each other. First, it fits into the economics of fertility. Historically, fertility rates have fallen with income and female labor force participation across countries and families. Economists stressed ideas like the quantity-quality trade-off of children (Becker, 1960; Becker and Lewis, 1973) and the role of the opportunity cost of women's time (Butz and Ward, 1979) to explain these empirical regularities.[4] However, neither of the aforementioned two relationships hold anymore in developed countries: the first one has flattened and the second one is now slightly positive. The new economics of fertility (Doepke et al., 2022b) stresses the importance of compatibility between career and family as a major driver of fertility. As a result, it sees it as crucial to consider these decisions jointly with labor force participation, as this paper does.

In the literature addressing these issues, many studies deal with them in isolation.[5] Moffitt (1984) and Hotz and Miller (1988) were some of the first to consider them jointly, in reduced form. Erosa et al. (2002) and Erosa et al. (2010) do so as well, but within structural models with exponential lives and in steady-state equilibrium. Francesconi (2002) is the pioneer study involving a dynamic life-cycle model with joint fertility and labor force participation decisions. Other papers that have featured such models since then include Da Rocha and Fuster (2006), Sheran (2007), Del Boca and Sauer (2009), Keane and Wolpin (2010), Erosa et al. (2016b), Adda

[4] For a survey on what Doepke et al. (2022b) call first-generation models of fertility, see Hotz et al. (1993)

[5] Early papers studying fertility in a static context include Becker (1960), Becker and Lewis (1973) and Willis (1973) Dynamic models of fertility that kept labor force participation constant include Ward and Butz (1980), Wolpin (1984), Rosenzweig and Schultz (1985), Cigno and Ermisch (1989), Heckman and Walker (1990), Blackburn et al. (1993), Arroyo and Zhang (1997) and Sommer (2016). Conversely, Heckman and Macurdy (1980), Blau and Robins (1988), Eckstein and Wolpin (1989), Van der Klaauw (1996), Hyslop (1999), Attanasio et al. (2008), Keane and Sauer (2009) and Blundell et al. (2016) study labor supply decisions in a dynamic framework while taking fertility decisions as exogenous.

et al. (2017), Guner et al. (2020) and Guner et al. (2021). Our main contribution to this literature is how we deal with heterogeneity in fertility preferences. We provide evidence that the desired number of children, as retrieved from fertility surveys, changes slowly across cohorts, does not heavily depend on income, and therefore provides a reasonable estimate of the lower bound for the potential fertility. Hence, we use desired fertility data to infer the distribution of preferences over the number of children. Moreover, efforts have been made recently to measure the impacts of maternity on labor market outcomes over different time spans, also known as child penalties (Kleven et al., 2019b). We use the event-study approach developed by these authors in our simulated data and contrast the resulting child penalties over time with the ones observed in the data.

This paper also contributes to the literature studying the impact of family policies on fertility and female labor force participation. There are broadly two branches in this body of research, one empirical and one structural (model-based). The empirical literature most often uses difference-in-difference and regression discontinuity designs to exploit natural experiments created by policy implementation. This way, they can identify the effects around it in a small window of time. In particular, Milligan (2005) and González (2013) study the impact of cash transfers in Canada and Spain, respectively.[6] The latter study estimates the short-run impact of this policy on fertility and mothers' labor supply that we use to calibrate our model. While they both find positive effects, the authors warn that the identified effects may include changes in timing, not just overall quantity. Cohen et al. (2013) use panel data on Israeli women and exploit variations in child subsidies to identify their effects on fertility. Interestingly, they find positive effects among older women, which they interpret as evidence in favor of an increase in the overall total of children and not just timing. In our results, these same effects explain the larger fertility effects in the short run.

On the other hand, the models used by the structural branch of the literature provide a laboratory to perform counterfactual quantitative experiments. This is important for three reasons. First, long-run impacts can be assessed by simulating the long run. Second, the results from the empirical literature strongly suggest that the context in which family policies are implemented matters (Cascio et al., 2015; Olivetti and Petrongolo, 2017). Thus, it is crucial to understand how they interact with one another and with labor market institutions. Policies can be implemented separately and jointly in the model, allowing researchers to understand better the interactions that arise. Third, it is possible to experiment with policies that have never been implemented, which means they cannot be tested empirically yet. Notice

[6] Other family policies have been studied as well. Bettendorf et al. (2015), Geyer et al. (2015), Givord and Marbot (2015), Haeck et al. (2015), Nollenberger and Rodríguez-Planas (2015) and Müller and Wrohlich (2020) evaluate the effects of childcare subsidies, while Lalive and Zweimüller (2009) and Asai (2015) study parental leaves.

that in this paper, we do all three things. Examples of this kind of work include some papers mentioned earlier: Erosa et al. (2010) study parental leaves, while Adda et al. (2017) and Guner et al. (2020) study cash transfers. Moreover, Bick (2016) studies the effects of childcare subsidies in a structural model featuring decisions over the total number of children and labor force participation but no decisions on the timing of fertility.

The main contribution of this paper is to connect the two branches of the literature studying the impacts of family policies on fertility and female labor force participation in the sense that we use a structural model, but we calibrate its parameters using carefully identified results from the empirical literature. We believe doing so is a powerful combination to understand better how family policies affect these outcomes. Indeed, there are direct appeals to this type of exercise in the literature, i.e.: "Combinations of clean designs with structural models of the sort presented in this paper may therefore be an avenue that helps to explore the longer-term effects of policy interventions" (Adda et al., 2017).

The remainder of the paper is organized as follows. Section 2.2 discusses the setting and lays out some descriptive evidence. Section 2.3 describes the model. In section 2.4, we explain how we take the model to the data and show the calibration results. Section 2.5 presents the main quantitative experiments. Section 2.6 concludes.

2.2 Background and Descriptive Facts

In this section, we describe the cash transfer policy that originated the natural experiment, which provides key moments we use as part of our calibration targets and the demographic and institutional context in which it happened. Then, we present evidence on fertility preferences that points to a gap between desired and realized fertility and discuss the robustness of these preferences to income and policy circumstances. Finally, we show evidence of changes in labor market behavior associated with motherhood. The facts presented in this section inform important modeling choices fully explained in section 2.3.

2.2.1 The Baby Check

Since the beginning of the 1990s, Spain has consistently been one of the countries with the lowest fertility rates in the world.[7] Against this backdrop, on July 3 2007, the Spanish prime minister unexpectedly announced the introduction of a one-time €2500 child benefit to be paid to mothers immediately after each birth. This benefit

[7] See appendix A2.1 for an overview of Spain's and similar countries' past and future demographics.

was complementary to any pre-existing child support or assistance. The policy came to be known in the Spanish media as "cheque bebé" (baby check, hereinafter). The projected population aging is explicitly mentioned in the originating law as the motive behind the new benefit, with the implicit understanding that it would increase fertility and slow down that demographic trend. The mechanisms through which the government expected the baby check to stimulate fertility were also made explicit in the law: compensating the additional expenditures households incur upon the birth of a new child and facilitating work-life balance. Lastly, the law mentions that the benefit was intended to maintain disposable income and welfare for low-income families with a new child.[8]

The baby check requirements were quite low, making it essentially universal. No income tests were stipulated, and the only real precondition was to have effectively resided in Spain for the two years before the birth[9]. This generated criticism to the benefit since its inception, as opposition parties argued that it was unfair that high-income families were receiving a similar benefit as low-income ones.[10] Nevertheless, the first baby checks were rolled out in November 2007. Over the next year, the government reported paying almost half a million of them, at the cost of approximately €1.2 billion, around 0.8% of the projected public expenditure for 2008.[11]

González (2013) exploits the sharp cut-off established by the government as a source of quasi-random assignment to identify the effects of the transfers on mothers' labor force participation.[12] She finds that mothers who received the benefit were 2-4 percentage points less likely to work twelve months after delivery. Moreover, she uses a differences-in-differences approach to identify the effect of the transfers on the annual number of births and finds that it increased by about 6 percent. These estimates provide worthy information about the short-term sensitivity of households' fertility and female labor force participation decisions to unearned income.

The baby check was phased out three years later, on May 12 in 2010, as part of overall budget cuts implemented by the Spanish government after the 2008 financial crisis.[13] The baby check was no longer available to women who gave birth after January 1, 2011. Similarly to the introduction of the child benefit, its cancellation was unexpected. González and Trommlerová (2023) quantify and compare the effects

[8] *LEY 35/2007*, November 15, 2007.

[9] Even this requirement was relaxed in 2009, making all women residing in Spain eligible for the benefit.

[10] "El Congreso aprueba el cheque-bebé". *El País*, October 18, 2007. https://elpais.com/elpais/2007/10/18/actualidad/1192695432_850215.html (in Spanish).

[11] "Presupuestos Generales del Estado, 2008". *Ministerio de Economía y Hacienda*, December 26, 2008.

[12] Note that the baby check was announced on July 3, and all mothers giving birth from July 1 on were eligible to receive it.

[13] Montero, Vicky, 2010. "José Luis Rodríguez Zapatero dice adiós al cheque-bebé, su medida estrella en política social". *RTVE*, May 5. https://www.rtve.es/noticias/20100512/331041.shtml (in Spanish)

of the introduction and cancellation of this policy. They find that right after the cancellation announcement in May 2010, there was a substantial rise in births of 4.7%, and right after the cancellation of the benefit, they observe a 5.7% decrease in births. They also find that the positive fertility response to the benefit's introduction in 2007 was smaller in magnitude than the fertility decline that occurred after the benefit's cancellation in 2010. Exploiting heterogeneity in economic conditions across provinces, they find that the larger negative effect of the cancellation was particularly pronounced in regions most affected by the economic crisis.

Researchers have quantified different family policies to boost fertility, but the issue remains unresolved in Spain and worldwide. On average, OECD countries devote slightly more than 2% of GDP on family benefits public spending, with Spain actually devoting only around 1% of GDP.[14] Moreover, a new baby check was introduced by the regional government of Madrid in January 2022, this time geared towards young mothers (below 30 years old), means-tested (families with annual income below €30000), and much more generous (up to €14500, in €500 installments).[15] The importance of assessing the effects of this kind of spending is therefore very much relevant.

2.2.2 Institutional Background

The impact of family policies is mediated by the context in which they are implemented. Here, we discuss the institutional background in which the baby check was introduced in July 2007 in Spain.

Public childcare provision at this time in Spain was pretty typical for a high-income country. Despite mandatory schooling starting only at age 6, preschool between 9 am and 5 pm was offered to all parents of children aged 3 to 6 that requested it, and take-up was nearly universal. For children aged 0-3, however, the availability of places in public childcare centers was quite limited, and parents, therefore, mostly had to pay for it out-of-pocket. Using the *Encuesta de Presupuestos Familiares* (Family Budget Survey, EPF henceforth) from the Spanish Institute of Statistics, we estimate that families with a child aged between 0 and 3 years old using full-time childcare spent on average €2000 a year in 2007.

The literature has pointed at taxation as one of the factors affecting the female labor supply. In particular, joint taxation has been identified as a strong disincentive to labor force participation among married women (LaLumia, 2008; Bick and Fuchs-Schündeln, 2017). Spain's system, however, allows couples to minimize their tax liabilities by filing jointly or separately, so it does not constitute an additional barrier

[14]Data from the OECD, "Family benefits public spending" in 2022. It includes mainly child-related cash transfers, parental leaves and income support for sole parents, and services for families with children.

[15]*Boletín Oficial de la Comunidad de Madrid núm. 308*, December 27, 2021.

to female labor participation.

One important aspect of the Spanish labor market that needs to be discussed here is duality, i.e., the coexistence of temporary and permanent contracts. The former have higher firing costs, and firms can effectively fire a temporary worker by not renewing her contract. Most young people start their careers in this situation, and with a probability that is initially low but increases with experience, they become workers with a permanent contract. This causes job insecurity, loss of earnings caused by intermittent unemployment, and lower returns to experience for young workers.[16] All these increase the cost of career interruptions early on for temporary workers, making it less likely that they can obtain a permanent contract. There is plenty of evidence that temporary contracts cause women to postpone and reduce their fertility (see Adserà (2004), De la Rica and Iza (2005), Auer and Danzer (2016), Lopes (2020) and Guner et al. (2021)). A related problem is the high unemployment rate. For example, Da Rocha and Fuster (2006) find that unemployment induces women to postpone and space births, which reduces fertility. However, the period before the policy implementation was one of low unemployment in Spain-the expansion before the Great Recession in 2008. Since this is the period we use for our model calibration, we do not model unemployment risk.[17]

2.2.3 Desired and Realized Fertility

In all high-income countries, women would like to have more children than they currently have, which may be one reason why policies such as the baby check might affect fertility. In this section, we use data from the *Encuesta de Fecundidad* (Fertility Survey, EdF henceforth) of the Spanish Statistics Institute to discuss the robustness of this fact for Spain.

There are two waves of the EdF, from 1999 and 2018. Crucially, women were asked in both of them about their desired fertility. The structure of the question, which is identical across years, is as follows: first, women declare the number of children they have. If they are childless, they are asked whether they would like or would have liked children, and if so, how many. If they have children, they are asked whether their number coincides with the number they would like or would have liked to have. If it does not, they are asked how many children they would like or would have liked to have. Finally, they are also asked to rank potential reasons behind this mismatch.

Table 2.1 shows the realized and desired fertility for women aged 40 to 44 in

[16] Dolado et al. (2021) argue that when the gap in firing costs is large between temporary and permanent contracts, as in Spain. Workers under the former exert less effort, and firms provide less training, which may explain this difference in the returns to experience.

[17] Bentolila et al. (2012) look at the role of duality in the unemployment surge in Spain during the Great Recession.

2018 and the difference between them. We chose this age group because most of these women had already completed their fertility, meaning the difference in desired fertility is definitive. The main takeaway from this table is that there are many women whose completed fertility is lower than their ideal. In particular, it shows that the fraction of childless women more than doubles the fraction of those who declare not wanting to have any children, while less than half the number of women that would have liked three or more children do. On average, the desired number of children for this cohort at this point in time was 2.04, while the actual number was 1.54, a gap of 0.5 children. This is very close to what is observed on average across OECD countries[18].

We believe that the existence of this gap is evidence that pronatalist policies have some margin to increase fertility. It can be reasonably argued that the answers given by people in surveys like the EdF do not represent policy-invariant preferences and that people respond by considering restrictions and incentives given by their current context. However, a pronatalist policy would very likely not diminish desired fertility by itself. Therefore, the gap between it and realized fertility can be seen as a conservative estimate of the maximum possible effect of such a policy. Nevertheless, in what follows, we present evidence that desired fertility would probably not change significantly in response to policies that alter decisions in the margin.

TABLE 2.1 Fraction of Women Aged 40-44 by Desired and Realized Number of Children

Number of children	Desired	Realized	Gap
0	7.90%	18.99%	-11.09%
1	15.20%	24.96%	-9.75%
2	49.75%	43.79%	5.95%
3 or more	27.15%	12.26%	14.89%

Source: *Encuesta de Fecundidad 2018, INE.*

An almost ideal way to test empirically whether the baby checks had any effect on desired fertility among the cohorts affected by it would be to use a Regression Discontinuity Design, similar to what González (2013) did for births and labor force participation. Unfortunately, we do not have data on desired fertility *right before* nor *right after* the policy announcement. We have two waves of the survey, which allows us to compare the same cohort's responses at different ages and cohorts at the same age.

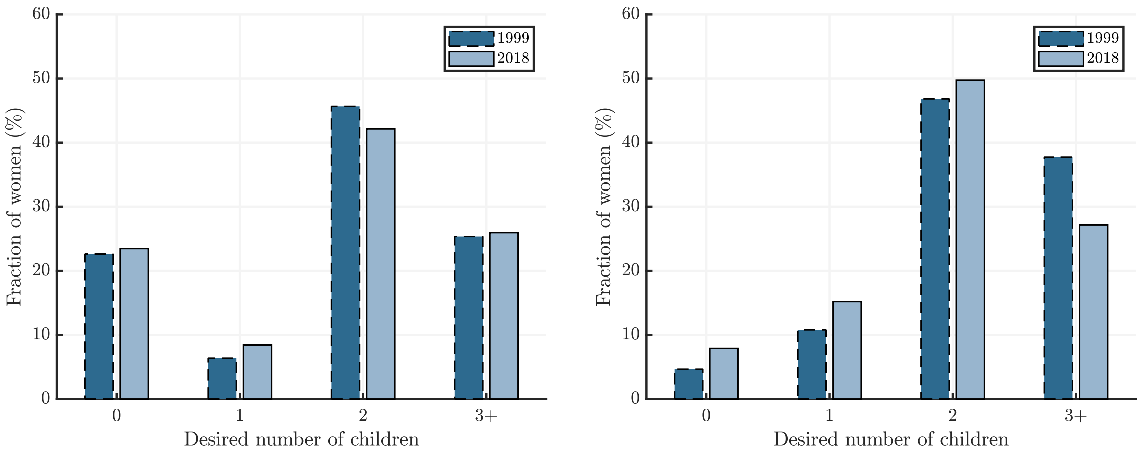
Figure 2.1 shows the desired number of children by cohort and by survey year. Consider first the cohort aged 21-25 in 1999 and 40-44 in 2018. When the baby check was introduced, these women were aged 29-33 in 2007 and therefore were affected by it. Their fertility preferences are represented by the dashed dark blue

[18]See OECD Family Database https://www.oecd.org/els/family/database.htm.

bar in Figure 2.1a and the solid light blue bar in Figure 2.1b. Notice that desired fertility experienced some changes between those years. In particular, the fraction of women that declared not wanting children in 1999 (when they were young, e.g., before most of them made any fertility decisions yet) is significantly smaller than the fraction in 2018 (after fertility was completed). Most of the difference is attributed to a higher proportion of women declaring they want one child. By contrast, the fractions wanting 2 or 3 children or more are similar. In other words, there is a discrepancy in the extensive margin but not so much in the intensive margin.

Two (non-mutually exclusive and non-exhaustive) hypotheses to explain this are that the baby checks changed the fertility preferences in a particular way for this cohort or that some women change their preferences as they age. The left panel shows that the desired fertility is almost identical across cohorts aged 21 to 25 and very similar across cohorts aged 40 to 44 in 1999 and 2018. This is evidence in favor of the second hypothesis. In a nutshell, Figure 2.1 points to fertility preferences changing somewhat along the life cycle but slowly across cohorts (which are exposed to different policies and economic conditions).

FIGURE 2.1 Fraction of Women by the Desired Number of Children for Selected Cohorts



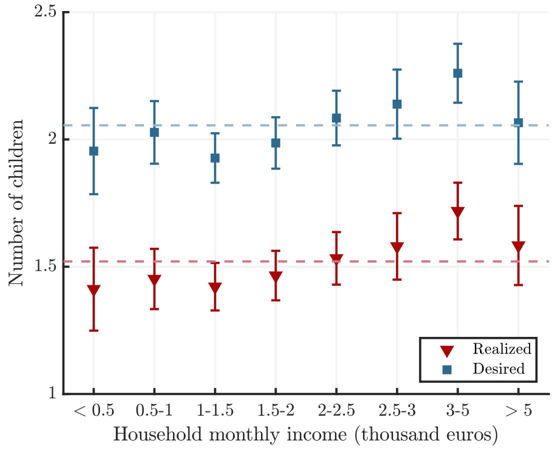
(a) Cohort aged 21-25 (b) Cohort aged 40-44

Source: Authors' work with data from *Encuesta de Fecundidad 1999 and 2018, INE.*

Moreover, Figure 2.2 shows the desired and realized fertility by household income for the same demographic as before. As discussed in the introduction, there's no clear relationship between fertility and income (if anything, there might be a slight positive correlation, whereas, in the past, it was negative). Moreover, desired fertility also does not exhibit a clear pattern. Since it does not change much with differences in family income shown here, which go up to an order of 10, it seems reasonable to conclude that desired fertility barely reacts to a one-time cash transfer which is a fraction of that.

Taken together, the evidence presented here indicates that fertility preferences reported in the EdF reflect, albeit imperfectly, deep parameters regarding the maximum number of children women would be willing to have under reasonable changes in policy.

FIGURE 2.2 Desired and Realized Fertility by Monthly Household Income, Women 40-44



Source: Authors' work with data from *Encuesta de Fecundidad 2018, INE.*
Note: Spikes represent 95% confidence intervals. Dashed lines represent the average realized and desired fertility of 1.52 and 2.05, respectively.

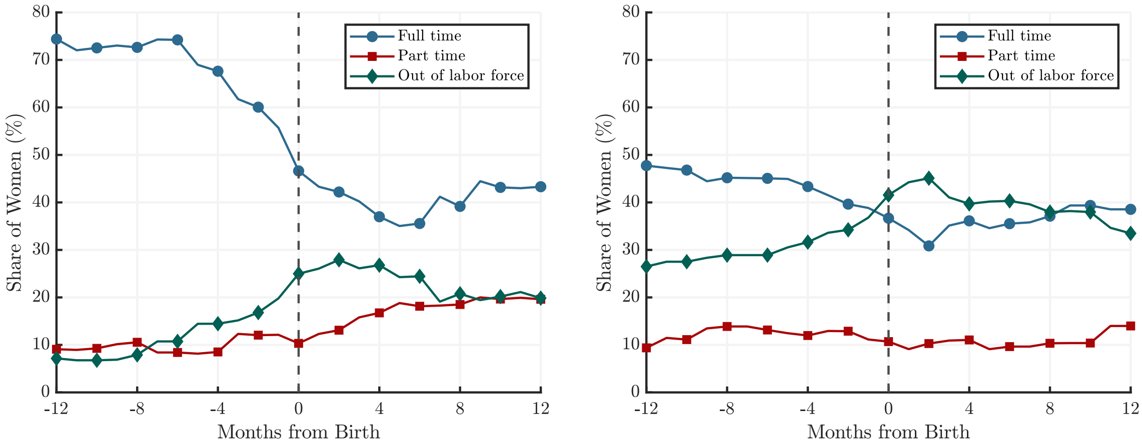
2.2.4 Labor Force Participation and Motherhood

Motherhood is associated with significant changes in women's labor market behavior. In this subsection, we use data from the *Encuesta de Condiciones de Vida* (Living Conditions Survey, ECV henceforth) of the Spanish Statistics Institute to provide some descriptive evidence for the magnitude of these changes among the cohorts of women affected by the baby check policy in the period immediately before its implementation. The ECV consists of a rotating panel, where households are interviewed for four consecutive years. There are 16000 households in the sample. Each year 4000 households leave, and 4000 new households replace them. We identified all births among women in the sample between 2004 (the first year of the sample) and July 2007 (that is, not eligible to receive the baby check). We kept all women for which there is labor force participation information for each of the 12 months preceding and succeeding the birth (24 months in total).

Figure 2.3 shows the fraction of women working full-time, part-time, and not in the labor force in this period by parity (first and second births). There is a large drop in full-time participation that actually starts about six months before the birth of the first child. The difference in the rate twelve months before and after childbirth is more than 20 p.p. Conversely, the part-time rate and the fraction of women that are out of the labor force increase by about 10 p.p. each[19]. By

[19]The fractions of women don't add up to 1 as some are also unemployed (not shown in the

Figure 2.3 Female Labor Force Participation Around the First and Second Births



(a) First birth (b) Second birth

Source: Author's work using *Encuesta de Condiciones de Vida 2005-2008, INE.*
Note: The data corresponds to the years 2004-2007.

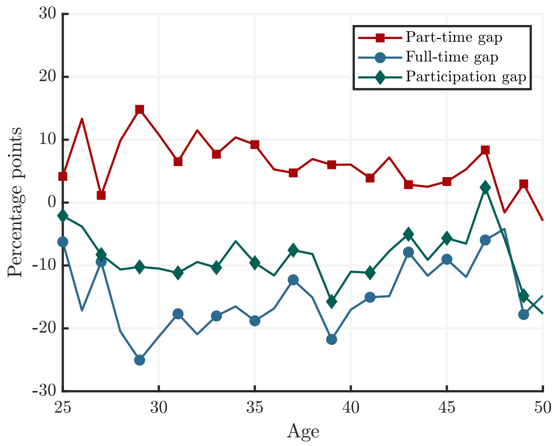
the time women give birth to their second child, full-time participation has barely changed compared to one year after their first one. However, the part-time rate has returned approximately to its pre-children level, while the fraction who are out of the labor force has inched upwards by between 5 and 10 p.p. The second childbirth is associated with another drop in full-time participation rates, this time of slightly less than 10 p.p. This is not compensated at all with part-time, and therefore almost entirely translates into a commensurate increase in non-participation over the next 24 months.

Taking stock, the first childbirth is associated with a significant drop in full-time participation, partially compensated by an increase in part-time. In contrast, the second birth is associated with a smaller drop that is not compensated by an increase in part-time and from a much lower baseline.

The previous analysis centers on the 24 months around childbirth and was performed on a balanced panel of women. To understand the longer-term effects of maternity on labor market behavior, we look at the cross-section of women present in the sample during the same period. Figure 2.4 plots the gap in participation, part-time and full-time rates between mothers and childless women by age. In general, mothers are much less likely to work full-time, somewhat more likely to work part-time, and less likely to work overall. The participation gaps close as women age and presumably their children grow up and start requiring less time, thus allowing them to work more easily. However, they persist between ages 30 and 40. Therefore, maternity is associated with a sudden drop in participation the year around childbirth and with lower participation rates for several years afterward.

graph).

Figure 2.4 Life-cycle Difference in Labor Force Participation Rates Between Mothers and Childless Women



Source: Author's work using *Encuesta de Condiciones de Vida 2005-2008, INE.*
Note: The data corresponds to years 2004-2007.

The main takeaway from the evidence presented here is that it is important to model the costs that children impose on mothers, especially time costs, in such a way that it generates the behavioral changes in the labor market associated with motherhood discussed above.

2.3 The Model

We study fertility and labor force participation decisions over the life cycle of married women. They derive utility from the number of children, consumption, and leisure. The model reflects the main trade-offs women face when making such decisions: each one has a target fertility level, but children negatively affect consumption and leisure time, so that desired fertility may not be achieved. The model captures dynamic career concerns as well. Becoming a mother at a younger and older age has various costs and benefits. Moreover, women face a trade-off between being a working mother or taking a costly career break once a baby is born. Labor market participation interruptions are costly because experience is not accumulated. Moreover, most women start working life under a temporary contract. To increase the chances of converting it to a permanent one, they must work to accumulate experience. However, there are costs of being a working mother. Considering all these constraints, each woman decides how many children to have and when to have them. The details of the model are covered in the rest of this section.

2.3.1 Demographics

We model women's life choices between the ages of 25 and 52. A model period corresponds to one year. We denote model periods with j, hence age 25 corresponds to $j = 1$ and age 52 to $j = J = 28$. Before $j = 1$, each woman is exogenously matched with a spouse/partner the same age as her and draws a desired number of children $N^* \in \{0, 1, 2, 3\}$. There are no marital transitions (no separations or divorces), and N^* remains constant for their whole life.

From ages 25 to 39, in each period, women must decide whether to try to have an additional child or not. We denote this decision by $b \in \{0, 1\}$. If she decides to do so ($b = 1$), she gives birth to a baby next period with probability α_j, which decreases with age. Therefore, women can give birth between the ages of 26 and 40. Each woman can have a maximum of three children, and we denote the total number of children by $N \in \{0, 1, 2, 3\}$.

After being born, each child transitions stochastically between four stages of life, based on the different kinds of costs children impose on mothers as they age. The first two stages are newborn (below one year old, thus born in the current period) and baby (between ages 1 and 3). The only difference between them is that newborns may come with a baby check. Children in these two stages entail the same consumption, leisure, and childcare costs. The third stage represents school-age (between ages 3 to 12). As children move to this stage, they continue to impose consumption and leisure costs, although quantitatively not identical to those imposed by newborns and babies. However, full-time childcare for school-age children is free for the parents (we assume universal public provision). Finally, those in the fourth stage are teenagers/young adults (ages 12 and up), which are not costly at all for mothers in terms of leisure but represent a larger burden on consumption. The exact mathematical form all these costs take is discussed in the following subsection.

We denote by $\mathbf{n} = [n_0, n_1, n_2, n_3]$ the vector indicating the number of newborns, babies, school-age children, and teenagers, respectively (hence $N = n_0+n_1+n_2+n_3$). A newborn becomes a baby with probability one after one period. A baby becomes a school-age child, and a school-age child becomes a teen with probabilities $\lambda_1 = 1/2$ and $\lambda_2 = 1/11$, respectively. In expectation, a child spends two years as a baby and eleven years of school age. The teenage/young adult stage is absorbing: once a child reaches this stage, it remains there. We denote this structure by $\mathbf{n'} = \Lambda_j(\mathbf{n}, b)$.[20] The main advantage of modeling children's aging in this way is that we avoid carrying each child's age as a state variable.

[20]See appendix A2.4 for the exact functional form this structure takes.

2.3.2 Preferences and Constraints

Women are the sole decision-makers in the household. Alternatively, they hold all the bargaining power, and thus the household's preferences are perfectly aligned with hers. In each period, apart from their fertility decision, they must also decide their labor force participation $h \in \{0, 1/4, 1/2\}$, where $h = 0$ means no participation, and $h = 1/4$ and $h = 1/2$ stand for part-time and full-time work, respectively.[21]

In each period, utility is derived from consumption (c), leisure (l), and children ($\mathbf{n}$). The functional form for instantaneous utility is given by the sum of CRRA terms for the first two, and an additional term which is a function of the latter:

$$u\left(c, l, \mathbf{n}; N^*\right) = \frac{\left(\frac{c}{\psi(\mathbf{n})}\right)^{1-\gamma_c} - 1}{1-\gamma_c} + \delta_l \frac{l^{1-\gamma_l} - 1}{1-\gamma_l} + \Gamma\left[\mathbf{n}, h; N^*\right], \tag{2.1}$$

where:

$$\Gamma\left[\mathbf{n}, h; N^*\right] = -\delta_N(\mathbf{n}; N^*) \frac{\exp(j-\gamma_N)}{1+\exp(j-\gamma_N)} |N - N^*| + \mathbb{1}_{\{N>0\}}\zeta + \mathbb{1}_{\{h=1/2\}}\kappa(\mathbf{n}). \tag{2.2}$$

Children affect utility directly and indirectly. The first two terms of the above expression depend purely on the number of children, and are, therefore, the direct effect. First, women experience a utility penalty from the difference between the current (N) and desired or target number of children (N^*). This can be interpreted as a craving for children among women that are still able to have them, and lifetime regret among those that cannot. Either way, it is the reason why women have children in the model.

The intensity of the penalty for not having the desired number of children depends on two terms. The first one, $\delta_N(\mathbf{n}; N^*)$ introduces a non-linearity:

$$\delta_N(\mathbf{n}; N^*) = \delta_{N1}\left[1 - \delta_{N2}\mathbb{1}_{\{N=2, N^*=3\}}\right], \tag{2.3}$$

i.e., we allow for the marginal disutility of not having a third desired child to be different from the marginal disutility of not having the first and the second.

The second one introduces age variation. In particular, it is a sigmoid function that increases with age and is asymptotic to 1. That is, younger women experience only a fraction of the utility penalty older women do, but this fraction increases with age. This can be interpreted in various ways. One is that it captures factors not included in the model that cause women to postpone fertility, such as housing and partner disposition. Another one is that it is a child-specific discount rate that

[21] Out of the 24 hours a day has, we assume 8 are used for sleeping and personal care, leaving 16 hours to be split between work, childcare, and leisure. A full-time job (8 hours per day), therefore, represents $\frac{1}{2}$ of the time endowment, while a part-time one (4 hours per day) represents $\frac{1}{4}$ of it.

decreases with age, as opportunities to have additional children diminish. Finally, there is a fixed (dis)utility of motherhood, ζ.

The functional forms described above allow the model to reflect aspects of the quantum and tempo of fertility. Anticipating the calibration, it should be apparent that ζ is crucial for the extensive margin of fertility (i.e. remaining childless or becoming a mother), δ_1 and δ_2 for the intensive margin (having 1, 2, or 3 children), and γ_N for the timing of births.

The third term in $\Gamma[\mathbf{n}, h; N^*]$, $\kappa(\mathbf{n})$ intends to capture the difficulties associated with working full-time while having kids. This includes schedule conflicts between full-time childcare and work, and disutility from not spending time with children.[22] We allow this cost to vary on the age of the youngest child:

$$\kappa(\mathbf{n}) = \begin{cases} \kappa_1 & \text{if } n_0 + n_1 > 0 \\ \kappa_2 & \text{if } n_0 + n_1 = 0 \text{ and } n_2 > 0 \\ 0 & \text{if } n_0 + n_1 + n_2 = 0, \end{cases} \tag{2.4}$$

that is, if there is a newborn or baby present in the household, the utility cost of working full time is κ_1, if there aren't any newborns or babies. If there are school-age children, it is κ_2, and if there aren't any children other than teenagers, the cost is zero.

Children are costly in terms of consumption (non-childcare), and older children are more so. To reflect this, we use the OECD equivalence scale to adjust per-capita consumption in the household. It assigns a value of 1 to the household head, 0.7 to each adult, and 0.5 to each child. Teenagers are counted as adults, hence:

$$\psi(\mathbf{n}) = 1.7 + 0.7n_3 + 0.5(n_0 + n_1 + n_2). \tag{2.5}$$

Leisure time is also negatively affected by children. We assume that children require a minimum amount of time from the mother that increases sub-linearly to reflect economies of scale in the childcare production function (e.g., it doesn't take double the amount of time to prepare food for two children than it takes to prepare food for one). Moreover, we allow the minimum time required by children to vary on the age of the youngest. In particular:

[22] Childcare usually goes from 9 am to 5 pm, but a full-time job may not correspond to that schedule. The fact that parents with young children must drop off and pick up their children at determined hours themselves or have someone else do it introduces an additional cost to working full-time. Moreover, Guner et al. (2021) report on the unusual organization of the workday in Spain, with long lunch breaks that create split-shift schedules. For example, 50% of workers are still at work at 6 pm in Spain, compared to 20% in the UK.

$$\xi(\mathbf{n}) = \begin{cases} \xi_1\sqrt{n_0+n_1+n_2} & \text{if } n_0+n_1>0 \\ \xi_2\sqrt{n_2} & \text{if } n_0+n_1=0 \text{ and } n_2>0 \\ 0 & \text{if } n_0+n_1+n_2=0, \end{cases} \tag{2.6}$$

e.g., if there are two children in the household, the cost can be different depending on the ages. If both are school-age, the leisure cost for the mother is $\xi_2\sqrt{2}$, but if one is a baby, the cost is $\xi_1\sqrt{2}$. This provides flexibility for the model to reflect the needs children of different ages may have. Notice once again that teenagers do not impose any leisure costs.

With this, we can define leisure time for the mother, which is the residual of a time endowment of one unit minus hours at work h minus time required by children:

$$l = 1 - h - \xi(\mathbf{n}). \tag{2.7}$$

Finally, we assume that women in the labor force with newborns and babies are required to purchase childcare time for the same amount of time she works. Resources available for consumption at the household level are therefore the household's net income I^{hh}_{net} minus childcare costs:

$$c = I^{hh}_{net} - \lambda 2h\left(n_1+n_2\right), \tag{2.8}$$

where λ is the cost of full-time childcare (we assume part-time childcare costs half as much).[23]

2.3.3 Income

We seek to model income in a way that captures income risk faced by households, accounts for labor market duality, is capable of reflecting the degree of (after-tax) income inequality (across households and genders, i.e. reflects the gender wage gap), and accounts for the returns of experience at different ages. To this end, we model spouses' (gross) income as depending on a couple of correlated, persistent stochastic shocks, the type of contract available (temporary or permanent), and accumulated experience.

The stochastic shocks for the woman and her partner are given by ϵ^f and ϵ^m, respectively. The household draws a couple of initial shocks $\left(\epsilon^f_1, \epsilon^m_1\right)$ from an exogenous joint distribution, that subsequently evolves over time following an AR(1) process:

[23]It is possible that childcare costs do not increase linearly with the number of children (because of sibling discounts, etc.). However, we were not able to find good sources to determine the shape of the function.

$$\begin{matrix} \epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\ \epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \end{matrix}, \quad \begin{bmatrix} \nu^f \\ \nu^m \end{bmatrix} \sim N\left(\begin{matrix} \mu^f = 0 \\ \mu^m = 0 \end{matrix}, \begin{bmatrix} \sigma^2_{\nu^f} & \rho \\ \rho & \sigma^2_{\nu^m} \end{bmatrix} \right). \tag{2.9}$$

Experience is accumulated over time by working. We assume the partners always work full time, and therefore their experience in period j is $j-1$, $\forall j \in \{1, \dots, J\}$. Women experience changes from one period to the next according to:

$$x' = \begin{cases} x & \text{if } h = 0 \\ x \text{ w.p. } 1 - \pi_x & \text{if } h = \frac{1}{4} \\ x + 1 \text{ w.p. } \pi_x & \text{if } h = \frac{1}{4} \\ x + 1 & \text{if } h = \frac{1}{2}, \end{cases} \tag{2.10}$$

that is, the experience remains constant if she does not work, increases by one year if she works full time, and increases by one year with probability π_z if she works part-time. Notice that if $\pi_x = \frac{1}{2}$, half a year of experience is accumulated with one year of part-time work. However, we allow for $\pi_x \neq \frac{1}{2}$, possibly lower, to reflect a penalty of working part-time on experience accumulation. We denote this structure as $x' = \Pi_x(x, h)$.

We model labor market duality only for women. They draw an initial contract type $z_1 \in \{0, 1\}$, where 1 denotes permanent contracts and 0 temporary ones. In subsequent periods the probability of having a permanent contract depends on experience, the type of contract available in the previous period, and whether she worked in the previous period:

$$z' = \begin{cases} 0 & \text{if } z = 0 \text{ and } h = 0 \\ 0 \text{ w.p. } 1 - \pi_z(x) & \text{if } z = 0 \text{ and } h > 0 \\ 1 \text{ w.p. } \pi_z(x) & \text{if } z = 0 \text{ and } h > 0 \\ 1 & \text{if } z = 1, \end{cases} \tag{2.11}$$

that is, if a woman has a temporary contract and works, she becomes permanent with probability $\pi_z(x)$ (which depends on her accumulated experience). If she has a temporary contract and does not work, she will still only have a temporary contract available for her next period. If she already has a permanent contract, she will also have one in the next period (it is an absorbing state). We denote this structure as $z' = \Pi_z(x, h, z)$.

Putting all the elements together, full-time (potential) log income for the woman is:

$$\ln\left(y^f\right) = \eta_0^f + \Delta\eta_0^f \mathbb{1}_{z=1} + \left(\eta_1^f + \Delta\eta_1^f \mathbb{1}_{\{z=1\}}\right) x + \left(\eta_2^f + \Delta\eta_2^f \mathbb{1}_{\{z=1\}}\right) x^2 + \epsilon^f, \tag{2.12}$$

that is, the type of contract changes the baseline income level and the returns to experience. In particular, if $\Delta\eta_0^f > 0$, $\Delta\eta_1^f > 0$, $\Delta\eta_2^f < 0$ and $\pi_z(x)$ is increasing in x, women have an additional reason to work as much as possible at the beginning of their career, i.e. to increase the likelihood of getting an open-ended contract, under which expected income is higher.

Log income for the husband is given by:

$$\ln(y^m) = \eta_0^m + \eta_1^m (j-1) + \eta_2^m (j-1)^2 + \epsilon^m. \tag{2.13}$$

The household's net income is the sum of the gross incomes of both partners minus tax liabilities:

$$I_{net}^{hh} = I\left(y^m, y^f, h\right) - T\left(y^m, y^f, h\right) = y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right), \tag{2.14}$$

where ϕ is an earnings penalty on part-time work, and $T(.)$ is a tax liability function.

2.3.4 Timing, States, Choice Variables, and Problem in Recursive Form

Upon entering the economy, women draw a fertility preference N^*, an initial contract type z_1, and initial income shocks for them and their partners $\left(\epsilon_1^f, \epsilon_1^m\right)$. Then, they enter period 1 with no experience and no children, i.e. $x_1 = 0$ and $\mathbf{n}_1 = [0,0,0,0]$.

From period 1 on, women observe their state vector $\left[\epsilon^f, \epsilon^m, z, x, \mathbf{n}\right]$, and choose their labor supply h and whether or not to try to have an additional child next period $b \in \{0,1\}$. This continues for every period until they reach 3 children or age 39 ($j = 15$). After this happens, they cannot have any more children, and they choose only labor force participation in each period.

The dynamic problem women solve in period j is given by:

$$
\begin{aligned}
& V_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \\
& \max_{\substack{h \in \{0, \frac{1}{4}, \frac{1}{2}\} \\ b \in B_j(\mathbf{n})}} u\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}\left[V_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z^\prime, x^\prime, \mathbf{n}^\prime; N^*\right) \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}\right] \\
& \text{s.t.} \\
& l = 1 - h - \xi(\mathbf{n}) \\
& c = I^{hh}_{net} - \lambda 2h\left(n_1 + n_2\right) \\
& I^{hh}_{net} = y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right) \\
& \ln\left(y^f\right) = \eta^f_0 + \Delta\eta^f_0 \mathbb{1}_{z=1} + \left(\eta^f_1 + \Delta\eta^f_1 \mathbb{1}_{\{z=1\}}\right) x + \left(\eta^f_2 + \Delta\eta^f_2 \mathbb{1}_{\{z=1\}}\right) x^2 + \epsilon^f \\
& \ln\left(y^m\right) = \eta^m_0 + \eta^m_1\left(j-1\right) + \eta^m_2\left(j-1\right)^2 + \epsilon^m \\
& \epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\
& \epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \\
& \mathbf{n}^\prime = \Lambda_j(\mathbf{n}, b) \\
& x^\prime = \Pi_x\left(x, h\right) \\
& z^\prime = \Pi_z\left(x, h, z\right),
\end{aligned}
\tag{2.15}
$$

where the choice set for the birth decision is defined as:

$$
B_j(\mathbf{n}) = \begin{cases} \{0, 1\} & \text{if } N < 3 \text{ and } j < 15 \\ \{0\} & \text{otherwise.} \end{cases}
\tag{2.16}
$$

2.4 Calibration

In this section, we specify the calibration results, discuss the identification we follow for the Method of Simulated Moments (MSM), and discuss the model validation for targeted and non-targeted moments.

We calibrate the model in two steps. First, we take a set of parameters from previous literature or estimate without solving the model yet. These include the parameters of the income process, the pregnancy success probabilities by age, the distribution of women over the number of desired children, and the cost of childcare. The second set of parameters is chosen jointly by targeting a set of data moments concerning average labor force participation rates for women without children and mothers with children of different ages, the tempo and quantum of fertility, and the short-run response to the sudden introduction of a cash transfer policy. We rely on the MSM for these.

2.4.1 Parameters Chosen Before Solving the Model

Income Process. To estimate the parameters governing the household's gross income, we turn to Spanish administrative data. In particular, we use the *Muestra Continua de Vidas Laborales* (MCVL, Continuous Working Life Sample), which is a 4% random sample of all affiliates to Social Security in Spain. The first step is to estimate the following Mincer regression separately for men, women with temporary contracts, and women with permanent contracts:

$$\ln(y_{it}) = \beta_0^s + \beta_1^s x_{it} + \beta_2^s (x_{it})^2 + \Theta_{it} + \epsilon_{it}, \tag{2.17}$$

where x_{it} denotes the experience for individual i at time t, Θ_{it} is a vector of controls, and $s = \{m, ft, fp\}$ stands for men, women under temporary and women under permanent contracts, respectively.[24] From the results of these regressions, we obtain the constant term and the returns to experience for the gross income equations for each of the three groups. We then take the residuals obtained from this estimation and regress them on their time lags at the individual level, to obtain the persistence parameters of the AR(1) process for the stochastic shocks, and the variance of the innovations. For the correlation coefficient between spousal shocks, we follow Hyslop (2001) and set a value of 0.25.[25] The results of this procedure are shown in Table 2.2.

For the numerical solution, we approximate the auto-regressive vector of stochastic shocks for the woman and her partner with a discrete-valued Markov chain using the method proposed by Tauchen (1986) and Tauchen and Hussey (1991). We use a 10 by 10 grid for the values of the shocks, where each point is calculated so that the income of the n-th point is the average income of the n-th decile for men and women at age 26. For the initial distribution of households over that grid, we identified all marriages without children and in which the woman was between 26 and 30 years old in the ECV between 2004 and 2007. Then, we computed hourly wages, and the deciles for each (separately for women and partners). Finally, we created a 10 by 10 matrix containing the fractions of couples by woman and partner decile.[26] This is the distribution from which the initial income shocks are drawn.

Probability of contract transition. To estimate the probability of transitioning from a temporary to a permanent contract as a function of experience, we run a probit regression where the left-hand side variable is a dummy when a woman is on a permanent contract, on experience, experience squared, age, and education. We then estimate the marginal effect of one additional year of experience for the average

[24] See the appendix for the exact specification.

[25] This was estimated for the United States. Replicating his work for Spain is beyond the scope of the paper.

[26] There is a small number of married, childless, non-working women in this age range, which we discard.

Table 2.2 Parametrization of the Income Process

	η_0^f	$\Delta\eta_0^f$	η_1^f	$\Delta\eta_1^f$	$\Delta\eta_2^f$	η_2^f	ϕ^f	$\sigma^2_{\nu^f}$
Women	6.348	0.806	0.048	-0.007	-0.003	0.001	0.906	0.184
	η_0^m		η_1^m		η_2^m		ϕ^m	$\sigma^2_{\nu^m}$
Men	7.093		0.046		-0.001		0.900	0.184

woman for each level of experience (from zero to 25 years). We use the MCVL to estimate this. In the Appendix, Table A2.4 provides the values of $\pi_x(x)$.

Experience accumulation. For simplicity, we assume that the accumulation of experience when working part-time is stochastic, with a probability of accumulating one year of experience $\pi_x = \frac{1}{2}$. Thus, the expected accumulation of experience of one additional year of part-time work is half a year.

Tax function. We use the tax funtion estimated for Spain by García-Miralles et al. (2019) to account for tax liabilities and credits. Total tax liabilities are given by $T\left(y^m, y^f, h\right) = \tau I\left(y^m, y^f, h\right)$, where the average tax rate τ takes the form:

$$\tau = \begin{cases} 0 & \text{if } I\left(y^m, y^f, h\right) < \tilde{I} \\ \max\left\{1 - \tau_0 \left(\frac{I\left(y^m, y^f, h\right)}{\bar{I}}\right)^{-\tau_1}, 0\right\} & \text{if } I\left(y^m, y^f, h\right) \geq \tilde{I}. \end{cases} \tag{2.18}$$

That is, households with a gross income below the threshold $\tilde{I}$ do not pay any taxes and the average tax rate increases with the ratio of household to average income $\bar{I}$. In particular, $\tilde{I} = 1404$, $\bar{I} = 3900$,$\tau_0 = 0.8823$ and $\tau_1 = 0.1224$.

Pregnancy probabilities. The probabilities of pregnancy success conditional on age α_j are estimated following Sommer (2016), who fits an exponential function to point estimates of infertility by age from the medical literature (Trussell and Wilson, 1985). They imply a success probability of 0.85 at age 25, which drops slowly to 0.77 at age 30, somewhat more rapidly to 0.65 at age 35, and then to 0.48 at age 39.[27]

Distribution of women over the number of desired children. For the distribution from which women draw N^*, we use the fractions of women by the desired number of children taken from the EdF reported in Table 2.1. We are aware that

[27] See appendix A2.5 for more details.

these may be responsive to policy. However, in section 2.2, we discuss why it is unlikely that one like the baby check has a large effect on these answers, based on the small observed differences across average responses among people with different incomes in the cross-section, and the similarity of average responses across cohorts over time. In any case, we believe these answers capture relevant information about the heterogeneity in preferences among women in the population of interest and are useful in a calibration exercise.

Childcare cost. Using the Spanish Family Expenditures Survey, we estimated that a family with a child aged 0-3 using full-time childcare spent in 2007 on average around €2000. Therefore, we use this number as the cost per child per period for full-time and half the amount as the cost for part-time childcare.

Discount factor. We use $\beta = 0.96$ (Kydland and Prescott, 1982).

2.4.2 Parameters Chosen Via the MSM

The remaining 12 parameters are calibrated by matching 12 moments from the data. The targets can be divided into three groups. The first one comprises labor force participation rates for women with and without children. The second one encompasses the average quantum and tempo of fertility. The third group involves the response to the cash transfers identified by González (2013).

Labor force participation. We include part and full-time participation rates for three groups of women (6 targets): childless, mothers whose youngest child is 0-3 (newborn or baby), and mothers whose youngest child is 3-12 (school-age). The data moments are taken from the 2004-2007 ECV. For each woman, we have information on participation by month. We compute an average yearly participation rate, counting each month worked full-time as 1, each month worked part-time as $\frac{1}{2}$, and dividing by 12. Following Bick (2016), we create a yearly participation status variable that falls into one of our three categories: if the average participation rate was above 0.75, we count the woman as having worked full-time that year, if it is between 0.75 and 0.25 we count her as having worked part-time, and if it falls below 0.25 we count them as being out of labor force. We target the average part-time and full-time yearly rates for childless women between the ages of 25 and 51. For mothers, we target the average rates for all women with children of the respective age.

Number and timing of births. To compute the fertility targets, we use the EdF. In particular, we computed the fraction of women aged 40-44 in 2018 with 0, 1, 2, and 3 children and the average age at which this group had their first child. We use this sample because these women were likely very close to their completed fertility, which is the outcome of interest in the long run. These are 4 targets in total.[28]

Baby check. Finally, González (2013) states in the conclusions of her paper that she finds that after one year of the introduction of the baby check, the number of births increased by 6 percent, and mothers were 2-4 percentage points less likely to work after the introduction of the baby check in Spain. We take these numbers as her preferred estimates and target a 6 percent increase in births and a 3 percentage points reduction in participation. These are the last two targets. The number of births refers to the crude birth rate, the annual number of births per 1000 population. This short-term cross-sectional measure calculates the number of births one year after the policy implementation. To compute the model counterpart, we proceed as follows. For each cohort of women alive at the time of the policy introduction, t, the government announces that they are entitled to a baby check - conditional on having a baby- and they expect this to last forever. Considering this information, women update their decisions regarding the number and timing of births. Finally, we compare the number of births born in the economy in the period $t+1$, under the baseline economy-no baby check- and after the introduction of the policy. Therefore, this measure of fertility does not reflect the completed fertility rate, which is why we label it a short-term effect. To compute the labor force participation response, we follow the same methodology.

2.4.3 Discussion

Although all parameters affect all model moments once we solve the model, some are more important than others for certain targets. The first two parameters, γ_c and γ_l govern how fast marginal utility from consumption and leisure falls, respectively. They are therefore important in determining how willing are women to substitute between consumption, leisure, and children, and play an important role in the intensive margin of LFP decisions (part-time versus full-time) and for the fertility and LFP responses to the baby checks. The parameter on the age-varying weight on the fertility gap regulates how early the craving for children starts in the women's life cycle, thus it is important for the average age at first birth. The weight on leisure δ_l plays an important role in the extensive margin of LFP decisions. For the distribution of women by number of children, δ_{N1}, δ_{N2} and ζ are crucial. The first one increases the penalty for not achieving the desired number of children, and therefore when it is higher every woman is more likely to be closer to her desired number of

[28]One of the fractions of women by the number of children is residual, and therefore there are only three targets for quantum and one for the tempo of fertility.

children. The second one diminishes the marginal penalty of not having a third child and is thus important for the fraction of women that end up having 3 children. The last one is an additional utility from just being a mom, and plays an important role in the extensive margin of fertility, that is, the decision between remaining childless or having children. The next four parameters, ξ_1, ξ_2, κ_1, and κ_2 are relevant for the LFP decisions of mothers with children at different ages. The first two affect the extensive margin, while the last two the intensive one (by making full-time work costly). Finally, the part-time earnings penalty ϕ evidently has a direct impact on the likelihood of part-time work.

2.4.4 Model Evaluation: Targeted Moments

Table 2.3 shows calibrated parameters, with a description of their role in the utility function. Consumption utility is almost logarithmic, with γ_c very close to 1. The penalty for not having a third child is 27% lower than the penalty for not having the first two, while there is an additional utility from the first child as given by the fixed utility from motherhood ζ being positive. This can be interpreted as the utility cost of remaining childless being larger than the cost of having children but fewer than desired, which seems reasonable. The time cost with younger children is larger, while the extra cost of full-time work is very similar independent of the age of the children. Finally, the model implies that working part-time entails a 20% penalty on hourly earnings compared to working full-time.

Table 2.3 Parameters Calibrated With the Method of Simulated Moments

Parameter	Description	Value
γ_c	Curvature of consumption	0.985
γ_l	Curvature of leisure	0.151
γ_N	Age-varying weight on fertility gap	26.500
δ_l	Weight on leisure	0.832
δ_{N1}	Base weight on fertility gap	0.364
δ_{N2}	Weight on fertility gap, 3rd child	0.270
ζ	Fixed utility of motherhood	0.115
ξ_1	Time cost, youngest child 0-3	0.349
ξ_2	Time cost, youngest child 3-12	0.211
κ_1	Cost of full-time work, youngest child 0-3	-0.030
κ_2	Cost of full-time work, youngest child 3-12	-0.040
ϕ	Part-time earnings penalty	0.200

Table 2.4 shows the model's outcomes versus the data targets. In general, we achieve our objective of reproducing closely the data targets with the calibrated model. In particular, it is important that we were able to get the correct magnitude of the effect of the baby checks on fertility. We fall slightly short on the magnitude

of the effect on LFP, but we are not too far from the lower range of 2 percentage points drop in González (2013). The effect is nevertheless small, and the model is qualitatively close.

Table 2.4 The Model vs. the Data, Baseline Calibration Targets

Moment	Model	Data	Difference
Labor force participation:			
Childless women:			
Part-time rate	0.190	0.194	-0.003
Full-time rate	0.732	0.717	0.015
Mothers, youngest child 0-3:			
Part-time rate	0.285	0.276	0.009
Full-time rate	0.539	0.537	0.002
Mothers, youngest child 3-12:			
Part-time rate	0.285	0.276	0.009
Full-time rate	0.579	0.583	-0.004
Fertility:			
Share of women with:			
0 children	0.190	0.190	0.000
1 child	0.240	0.250	-0.009
2 children	0.444	0.438	0.006
3 children	0.126	0.123	0.004
Average age at first birth	29.458	29.300	0.158
Effects of cash transfers on:			
Annual number of births	0.062	0.060	0.002
Mother's LFP over the first year	-0.016	-0.030	0.014

Note: The effect of the cash transfers is the number of births over the next year after the introduction of the policy. The effect on mothers' LFP over the first year is the difference between participation rates. Data points are the ones reported by González (2013) in the conclusions of her paper, which we take to be her preferred ones: 6 percent increase in the annual number of births, mothers 2-4 percentage points less likely to be working 12 months later.

2.4.5 Model Evaluation: Non-targeted Moments

Here we discuss the model's results for a set of non-targeted moments. First, we would like to know how well the model replicates the long-term costs that motherhood imposes on women. Kleven et al. (2019b) propose an event-study specification around the birth of the first child to measure the effect of children on the labor market outcomes of the parents. This specification, originally applied to Danish data, has since been used by Kleven et al. (2019a) for Sweden, Germany, Austria, the UK, and the US. More importantly for us, De Quinto et al. (2021) did it for Spain, using the MCVL (which is the same data that we use to estimate the parameters of our income process). They assess the impact of the first child on gross earnings, days of work, probability of part-time employment, and probability of being on a temporary contract. The first and the last two of these outcomes have direct counterparts in our model. We implement the event-study specification on our simulated data. The results, together with the original estimates by De Quinto et al. (2021), are shown in Figure 2.5.

Qualitatively, the model displays adequate behavior in all three outcomes: the earnings and part-time penalties increase rapidly in the first two years after giving birth to the first child, while the temporary penalty is close to zero at the beginning and then increases slowly with age. Moreover, the three outcomes show a long-term penalty, i.e. even 10 years after giving birth the gap is still present.

The reason why earnings and part-time probability increase immediately after the first birth is straightforward: after having their first child, many women switch to part-time employment or drop out of the labor force, which means they earn less or nothing at all (the earnings child penalty is estimated unconditional on employment status). The reason behind the shape of the penalty on the probability of being on a temporary contract is more subtle: women on permanent contracts are more likely to have children (in the data you can see it as a negative penalty for the first two years, in the model, there is only a very small effect the period the child is born). However, over time those that had children but were on temporary contracts convert them at a much lower rate than those who didn't have children, because many of the former take career breaks or switch to part-time.

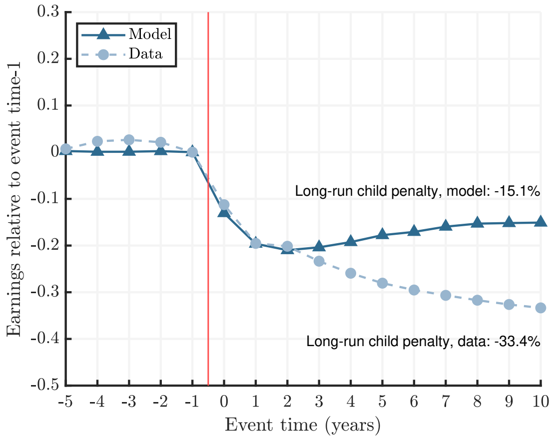
Quantitatively, the model's long-term part-time and temporary contract probability penalties are quite close to the data counterparts. The main discrepancies occur in the first 5 years of the part-time and the last 6 of the earnings penalties. We think this is due to the fact that there are a number of factors that the model does not feature that affect earnings among mothers, including occupational choice and loss of skills (Adda et al., 2017). During the first years after birth, the model gets the right earnings by overestimating the part-time penalty. After the fifth year, the part-time penalty falls very close to the one from the data. The earnings penalty reflects it by diverging from the one from the data.

While we target average labor force participation by childless women and women with children of different ages, we would also like to know how well the model reproduces the participation gaps between childless women and mothers at different ages. Figure 2.6 shows this, along with the data counterparts (which are the same series we showed in Figure 2.4). The model reflects very well the general patterns: overall negative participation and full-time gaps and positive part-time gaps that are the largest around age 30 and close gradually as women age.

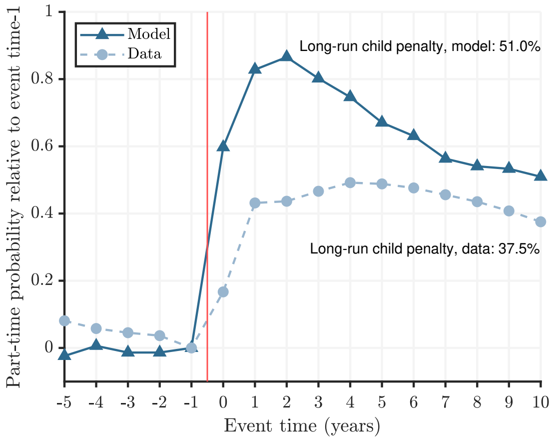
Finally, although we target the average age at first birth, it is interesting to

Figure 2.5 Child Penalties in the Model and the Data

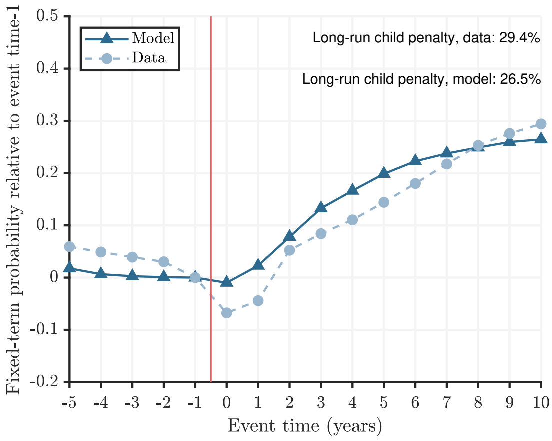
(a) Earnings



(b) Part-time work



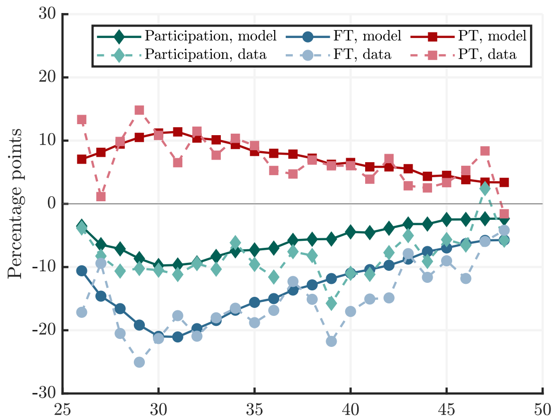
(c) Fixed-term contract



Source: Source: Author's work and data from De Quinto et al. (2021).
Note: The effects on (gross) earnings are estimated unconditionally on employment status. The effects on part-time and fixed-term contracts are estimated conditional on working.

compare how its distribution looks with respect to the data. This is shown in Figure 2.7. Again, qualitatively the model reproduces the main features: the fraction of women having their first child is positive for every age between 26 and 40, with

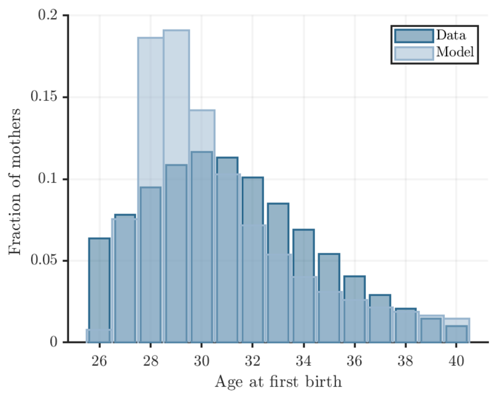
Figure 2.6 Life-cycle Labor Force Participation Gap Between Mothers and Childless Women in the Model



Source: Author's work and data from *Encuesta de Condiciones de Vida, 2004-2007, INE.*

the bulk of women having it in their late 20s and early 30s, and the fraction of first-time mothers decreasing slowly after that. The main discrepancies occur with women aged 26 and with women in their mid-thirties. However, the model displays a reasonable amount of time variability in this dimension.

Figure 2.7 Distribution of Mothers by the Age at First Birth, Model vs. Data



Source: *Encuesta de Fecundidad 2018, INE.*

Overall, the model's fit along non-targeted moments is satisfactory, and it seems to be an adequate setting to perform the quantitative experiments necessary to answer the main research questions considered in this paper.

2.5 Quantitative Experiments

One of the main advantages of having a structural model like the one we propose in this paper is that it allows us to perform quantitative experiments consisting of counterfactual simulations. In this section, we present the results of three such experiments. For the first one, we assess the long-run effects of cash transfers on fertility, by simulating the life cycle of women eligible for the policy for the entire duration of their lives. Then, we analyze the effects of an alternative policy, consisting of subsidizing childcare for mothers of children aged 0-3, which is the one group for which there was no universal public coverage when the baby check was introduced. Finally, we explore the duality of the Spanish labor markets' role in fertility, its interplay with labor force participation, and the effects of cash transfers.

2.5.1 Short and Long-run Effects of the Cash Transfer

So far, we have replicated the short-run effect of the baby check found by González (2013). Whereas she finds that this cash transfer increased the number of births after the policy implementation, the long-run effect remains to be seen. It could be that the increase in the crude birth rate is a tempo effect - women are anticipating fertility-or a quantum effect-women are having, on average, more children. In this section, we aim to disentangle both and understand whether the baby check introduced in Spain is cost-effective, as this answer depends on whether women have, on average more children. We are therefore interested in the *completed fertility rate* (CFR, the average total number of children women have) in the long run. To obtain this number, we simulate women's life-cycle fertility and labor force participation decisions using our model, with everyone eligible to receive the baby check in every period of their life and having full knowledge of the fact. Then, we can compare the CFR of a cohort of women in such a situation with the CFR in the baseline scenario with no cash transfers.

In our model baseline, the CFR was 1.553 children per woman. In our counterfactual exercise in which each woman is eligible to receive the baby check every period, the CFR goes up to 1.599. This is a 2.95% difference, just short of half of the 6% effect in the short run. This result leads to two main conclusions. First, the baby check has both tempo and quantum effects. Second, if we looked only at the crude birth rate, we would overestimate the policy's effectiveness. Therefore, to fully understand the cost-effectiveness of policies that target boosting fertility, it is imperative to look at the long run.

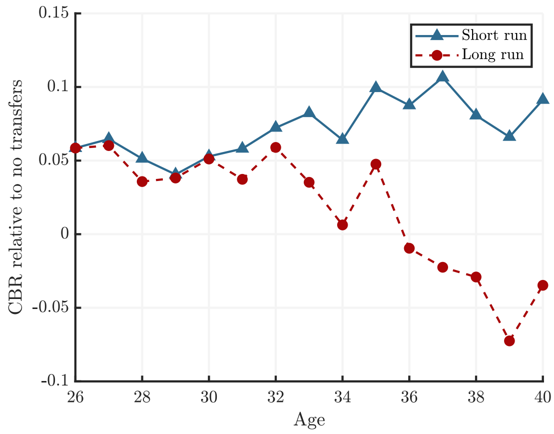
To understand why the long-run effects are different from those in the short run, it is helpful to consider the effect of the policy on women of different ages and to think about children as durable goods. After a certain age (around 30), most women want to have them as soon as possible (because they provide a utility flow). The exact timing depends on individual conditions. However, when a new policy makes it easier for women (at any age) to have children, the likelihood of them having a child in the following period increases for everyone. Moreover, for younger women, the likelihood of having them later may decrease. The profile of the crude birth rate (CBR, the fraction of women that have a baby in a certain period) by age, therefore, may look very different for a cross-section of women at the time the policy is introduced compared to a cross-section of women a few years later. The effect of the baby check on fertility found by González (2013) is on the total number of births in a short time window after the policy was implemented.

Figure 2.8 compares the effect of the cash transfer on the CBR by age in the short and the long run. The short-run effect is the CBR by age of the cross-section of women living at the time of the policy relative to the baseline CBR by age. That is, it shows the average fertility response for women of every age between 26 and 40 in the period following the policy announcement. The long-run effect is the CBR by age for women that were eligible for the baby check for all of their reproductive lives, i.e., for ages 25 to 39. Consider the cohort of women aged 30. In the short run, the CBR compares the number of births among 30-year-old women who receive the baby checks in the current year (at age 30), assuming such entitlement will continue until age 40, to the number of births among 30-year-old women in the baseline economy without baby checks. On the other hand, the long-run CBR measures the difference in the number of births at age 30 for women who had access to baby checks since their entry into the economy, compared to the baseline economy where such benefits were unavailable.

The short and long-run CBR are identical for the youngest women. This is because these women face identical horizons. The two lines remain close for a few periods but start to diverge in the 30s. The short-run effect remains positive and has a soft U-shape. The reason is that cash transfers have a larger effect on young women, who are poorer, and on older women, for whom the transfer represents an incentive to take one of the few last opportunities to close the gap between their realized and desired number of children. Women in their late 20s and early 30s were having children anyways, so the effect on them is smaller. The long-run effect starts to fall and even becomes negative in the late 30s. That is because young women had time to plan the births of the children they wanted before, and there are no older women who are caught by surprise by the policy anymore.

To further understand how the effect of cash transfers on births changes over time, we simulate fifteen cohorts of women of different ages when the baby checks are announced. In year -1, the policy is announced. Then, in year 0, the overall crude birth rate depends on the responses of women aged 25-39 that were caught by surprise by the policy in the previous period and changed their decisions accordingly. In year 0, women aged 26-39 make fertility decisions conditional on their previous period choices, which affect the next period. The average decisions of each cohort in period 0 may differ from those of the previous cohort in period -1 (when they were the same age). This is because some women in the newer cohort had an additional child, and that lowers the likelihood that they will have one during this period. Figure 2.9 shows the crude birth rate by year after the implementation of the policy.

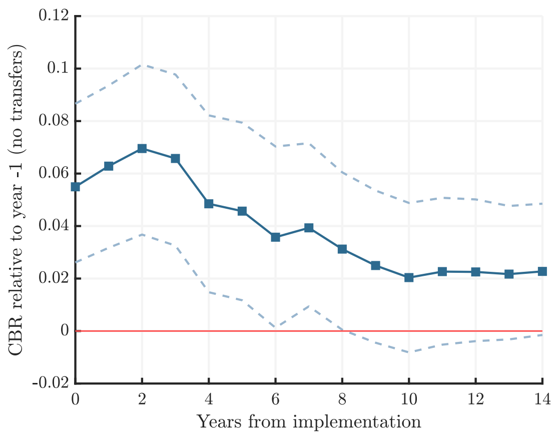
Figure 2.8 Effect of the Cash Transfers on the Crude Birth Rate (CBR) by Age



Note: The short-run effect of the cash transfers is the CBR by age for the cross-section of women living at the time of the implementation of the policy, relative to the baseline (without cash transfers). The long-run effect is the CBR by age among women that had access to the policy in every period in which they could have children, relative to the baseline.

The CBR in the period right after the implementation goes up by around 6%, but then it gradually falls to around 3%, as the CBR profile by age converges to the long-run one shown in Figure 2.8.

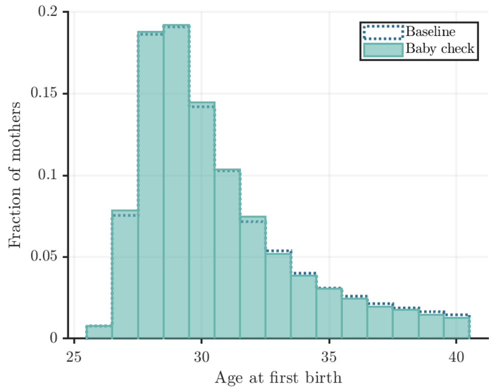
Figure 2.9 Overall Crude Birth Rate Relative to no Cash Transfers by Year After Implementation



Note: Dashed blue lines are 95% confidence intervals, constructed via bootstrapping.

By observing the shape of the long-run effects of cash transfers by age one concludes that timing must also change. In Figure 2.10 we have plotted the fraction of women by age at first birth in the baseline and in our long-run counterfactual exercise. As expected, the distribution under the policy has more mass at younger and lower mass at older ages. In particular, there are more first-time mothers in their late 20s and fewer in their late 30s when cash transfers are in place. The average age at first birth is reduced with the transfers by 0.15 years (about two months), while the average spacing between children is also shortened, but the magnitude of the change is smaller (less than a month between the first and second and slightly more between second and third).

Figure 2.10 Age at First Birth in the Model, Baseline and With Baby Check



In a nutshell, the main reason why the short and long-run impacts of the cash transfers are different boils down to the difference between the fertility response that women in different stages of the life-cycle have to the surprise announcement and the response when they have had time to adjust the timing of their fertility. For younger women, that response is not that different. For older ones, it is, because they cannot go back in time and change their previous decisions and have only a few periods to adjust. Over time, this additional effect from older women having to adjust immediately goes away, as the women that become old have had time to adjust.

Discussion

Our results are consistent with the discussion in González and Trommlerová (2023). They find evidence for a quantum effect, a rise in cohort birth rates. First, the baby check increases fertility regardless of parity. It would be unlikely to see a rise in fertility among women who have already had two children if the fertility increase were only a tempo effect. Second, the baby check increases birth rates among older women. Despite this evidence, the short duration of the policy (3.5 years) and the economic crisis surrounding the baby check, make evaluating its long-run effects empirically challenging. We contribute to this paper by evaluating the long-run effect of this policy. We find both a tempo and quantum effect, although the latter is smaller than the rise in the crude birth rate found after the policy implementation-the short-run effect.

2.5.2 Childcare Subsidies

Childcare availability (or lack thereof) is frequently cited as one of the reasons why fertility may be low. Subsidizing childcare is a natural alternative to a universal cash transfer, such as the baby check. Here, we consider the effects of a subsidy for children aged 0-3, since schooling after that age is provided for free in Spain.

To be comparable in magnitude to the payment offered by the baby checks, we compute the proportional subsidy with a present discounted value equal to the baby check payment, i.e., €2500, for a full-time working mother. Such proportional subsidy is 43.34% of the yearly cost of full-time childcare (€2000), for the first three years of an infant's life. Since, in our model, women must buy childcare for the amount of time they work, an alternative way of seeing this policy is a cash transfer conditional on working.

Not surprisingly, the long-run fertility effects are smaller than the ones observed with the baby checks. Completed fertility increases by only 2.06%, compared with 3% with cash transfers. The two policies diverge in their effects on labor force participation: it drops by a little bit more than 1 percentage point with the cash transfers with respect to the baseline, but it increases by almost 1 percentage point with the childcare subsidies. This result, again, is hardly surprising since the latter is essentially a subsidy to working mothers. Moreover, part-time work remains unchanged, meaning that all of the effects operate through full-time.

Finally, regarding the effect of the childcare subsidies on timing, again, they are slightly more muted than with the cash transfers. The average age at first birth decreases by close to a month and a half, while the spacing between children remains unchanged.

2.5.3 Impact of Labor Market Duality

Temporary contracts are associated with delayed and depressed fertility. While most of the literature exploring the effect of this type of working conditions on fertility centers on uncertainty and stability, we highlight another mechanism: returns to experience at critical ages. Most women start their careers under a temporary contract and anticipate their labor market participation may be reduced after having their first child. Therefore, they have an additional incentive to work full-time, accumulate experience and postpone childbirth, i.e., obtain a permanent contract first.

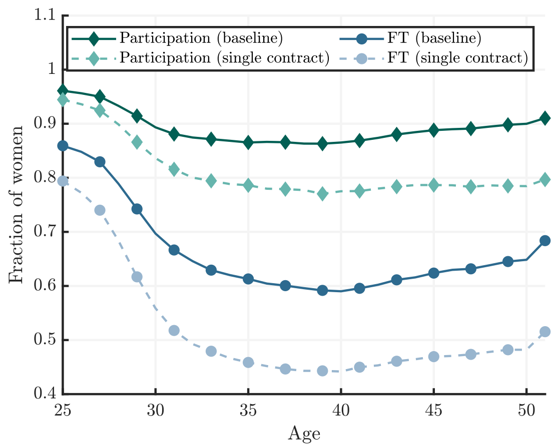
In our model, the difference between temporary and permanent contracts is reflected in the parameters of the income process. In particular, the constant term is larger for the permanent one, reflecting the fact that temporary workers experience unemployment spells during the year. Moreover, the returns to experience are larger for permanent workers, which could reflect firm or tenure-specific skill accumulation that temporary workers do not accrue.

The counterfactual we propose here is to have a single contract represented by a unique income process with parameters estimated from pooling all female workers. Among younger workers, this contract offers full-time earnings that are lower than the permanent contract, but higher than the temporary one in the baseline scenario. Therefore, the returns to experience are not as large at the beginning of women's careers.

We re-compute the model solution under the new parametrization for the income process, simulate the life cycle for cohorts of women that live under these labor market conditions, and compare it to the baseline scenario (with dual labor markets and no cash transfers). Our main result is a 9.6% increase in completed fertility rates, which is three times as much as the long-run increase we found with cash transfers. Moreover, the first age at birth is anticipated by 5 months, and the timing between births is slightly shortened.

To understand better what drives these effects, Figure 2.11 shows the average overall and full-time labor force participation in the baseline and under the single contract. The former is not very different early in women's working lives, but there is a gap in full-time participation early on. This gap widens as women get children (which happens earlier in the single contract scenario).

Figure 2.11 Female Labor Force Participation Through the Life Cycle, baseline, and single contract



Finally, the last experiment we carry out is to implement the cash transfers on top of the single contract and repeat the simulations. The completed fertility rate increases by 2.6% with respect to the scenario with a single contract but no transfers. This is very close but slightly lower than the transfers' effect on the economy with dual labor markets. In terms of timing, the average age at first birth is further reduced by 0.12 years, or about a month and a half, with further smaller reductions in spacing, again very close to the effect on timing on the economy with temporary and permanent contracts. These results imply that the results of our first experiment, i.e., the effects on fertility of giving cash transfers upon birth are smaller in the long run than in the short run and modest in magnitude, were not heavily dependent on that feature of the model and the Spanish economy.

2.5.4 Discussion: Bargaining and the Role of Fathers

One important assumption we make in our model is that women are the sole decision-makers in the household. Moreover, the explicit behavior of fathers is not analyzed. Here we discuss the potential implications of accounting for bargaining and allowing for a more active role for fathers.

In heterosexual partnerships like the ones considered in this paper, both sides must participate in the making of a baby, and therefore there needs to be some sort of agreement for it to occur. This decision has two dimensions: the overall number of children and the timing of births. In a standard bargaining framework, both preferences enter as inputs, and the resulting outcomes should be a sort of weighted average of the preferences of the husband and the wife, with the weights depending on the relative threat points and bargaining power. Indeed, that is what the literature finds. In developing countries, the disagreement over the total desired number of children is larger, with men preferring to have more of them and their preferences taking precedence, likely because they tend to have more bargaining power in these contexts.[29] In industrialized countries, women's preferences are, at the very least, as important as men's, and each partner enjoys veto power when deciding whether to have additional children.[30] Doepke and Kindermann (2019) account for this using a quantitative model of household bargaining, and conclude that fertility responds highly to interventions that lower the childcare burden of women.

Accounting for bargaining may have some implications for our results. While there is no particular reason to think cash transfers lower the burden of women, childcare subsidies may. Therefore, this latter policy may have additional effects to the ones found by our model. Moreover, there is evidence that the disagreement over the desired overall number of children is small in developed countries. In particular, we find very similar preferences in this respect in Spain.[31]This suggests that

[29]Westoff (2010) surveys desired fertility among men and women in African countries and finds differences of up to 5 children. Bankole (1995) and Gipson and Hindin (2009) provide evidence on the relative importance of men's fertility preferences in Nigeria and Bangladesh, respectively. Rasul (2008) develops and tests a model of household bargaining over fertility using Malaysian data. He finds that couples bargain without commitment and that fertility outcomes depend on the relative threat points, which vary across ethnic groups.

[30]Thomson (1997), Thomson and Hoem (1998), Testa et al. (2014) and Hener (2015) find evidence on this using data from the United States, Sweden, Italy, and Germany, respectively

[31]See the appendix for details.

using women's total desired number of children is not unreasonable in our context. However, the scope for disagreement on the timing of births may be important. As a rejoinder, two features of the model could be represented as a reduced form for the veto power that husbands may have on the timing of births: the age-dependent weight on utility from children and the taste shock we add for the computational solution.[32] Naturally, there may be interactions between policies like the cash transfers here and men's willingness to have children at different stages of life that the model, unfortunately, is unable to capture.

Whether or not one considers bargaining, it seems evident that fathers' behavior can affect the effect of family policies on fertility. The presence of a more cooperative partner, who, for example, is more willing and able to pick up kids from school, attend school meetings or cook meals, would relax the time costs children impose on mothers and their willingness to have more of them. Some of these effects could be similar to the effects of lowering the values of the parameters ξ_1 and ξ_2, which govern the time cost children of different ages impose on mothers and on κ_1 and κ_2, which represent the costs of being a full-time working mother. As an additional exercise, we lower by 5% the values of these parameters to assess how much this affects fertility decisions. The completed fertility rate increases by 10% in response to these changes. However, the effect of cash transfers becomes more muted. Partly this is due to fewer women being away from their desired fertility levels. There is no good reason to think that cash transfers would increase the father's willingness to do childcare and housework, and their modest effects on fertility are consistent with this. Other policies, like paternal leaves, may have this effect.[33]

2.6 Conclusions

A natural experiment involving cash transfers upon birth took place in Spain in 2007. The experiment was exploited by González (2013) via a DiD-RDD design to estimate its effects on fertility and mothers' labor force participation. The causal relationship identified thus is best interpreted as short-run, around the policy intervention.

In this paper, we develop a life-cycle model of fertility and labor force participation, and calibrate its parameters using these cleanly identified short-run effects, along with other aggregate moments for Spain in the period right before the implementation of the policy. Using the model, we explore the longer-run effects of the transfers, as well as the effects of alternative policy interventions, and their interactions with an important feature of the Spanish labor markets: the coexistence of

[32]See the computational appendix for more details.

[33]Ekberg et al. (2013) find no effect on parental leave-taking on household work, but only use as a measure for it the share of leave taken for care of sick children. Bünning (2015) and Tamm (2019) do find long-lasting effects of leave-taking on childcare and housework. Farré and González (2019) interestingly find that paternity leave reduces fertility using data from Spain.

temporary and permanent contracts (duality).

We find that the long-run effect of cash transfers on fertility is about half as large as it is in the short run. The main reason is that in the short run, there are additional births by older women that do not have time to adapt their previous fertility choices and have to adjust soon. Moreover, we find that a childcare subsidization policy that gives women the same amount in present value as the cash transfers do brings about an increase in long-run fertility that is only slightly lower, but increases labor force participation instead of decreasing it. Furthermore, we find that the duality in the Spanish labor markets has a large effect on fertility, about three times as big as the effect of the transfers, driven by an increase in returns to experience during crucial years for child-rearing. However, labor market duality does not seem to be driving our results regarding the short versus long-run effects of cash transfers on fertility, as evidenced by the fact that the results change very little when we implement the policy in a scenario where only a single contract is available.

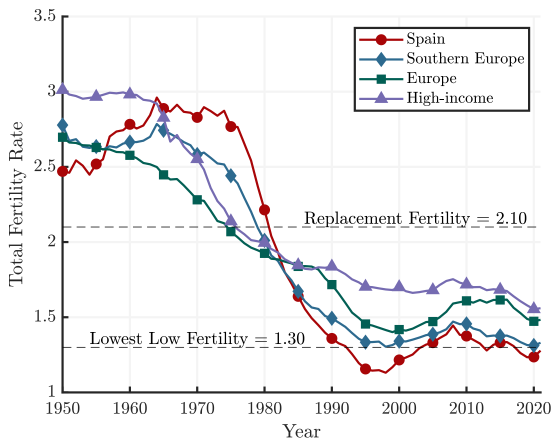
Our results highlight the importance of policy interventions that make labor market interruptions less costly (elimination of duality) and less necessary (childcare subsidies). This is in line with the main idea of the new economics of fertility, i.e. a crucial driver of it nowadays is the compatibility between career and family (Doepke et al., 2022b).

A2 Appendix to Chapter 2

A2.1 Demographics in Spain

In the early 1990s, Spain became one of the first countries in the world to attain what demographers call lowest-low fertility, i.e. a total fertility rate (TFR, from now onwards) below 1.3 (Kohler et al., 2002). A group of countries, mostly in Southern, Central, and Eastern Europe, followed. Figure A2.1 shows how Spain experienced a relatively late baby boom in the 1960s and 1970s, during which its TFR increased above that of others in its geographic proximity and that of the rest of high-income countries[34]. However, it fell rapidly in the 1980s, dipping below the lowest-low fertility threshold for a good deal of the three decades between 1990 and 2022, and below that of its peers for the entirety of that period.

Figure A2.1 Total Fertility Rate, Selected Countries (1950-2022)



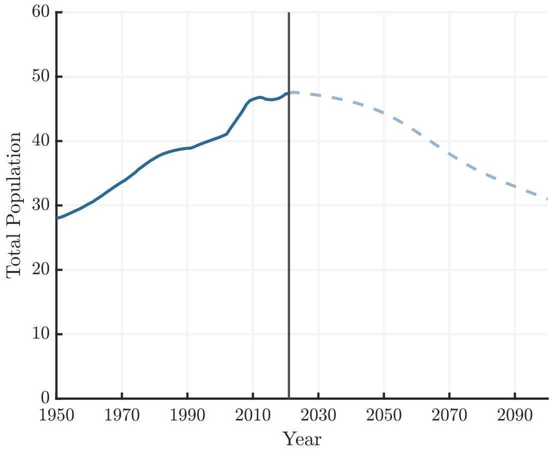
Source: United Nations, Department of Economic and Social Affairs, Population Division (2022).

Moreover, since the early 1980s, most of the developed world has had a TFR below the replacement rate of 2.1. This means, in the absence of immigration, that population will eventually decline. Figure A2.2 shows the estimated and projected population in the countries for the same group of countries. In all of them, the population peak is likely very close or has already passed. In the case of Spain, it is apparent that the population would have already peaked if it were not for the large influx of immigrants (mainly from Latin America) it received in the 2000s. In fact, considering the very low fertility levels, in the absence of large-scale immigration, it is all but certain that most of the countries considered here will see population decline soon.

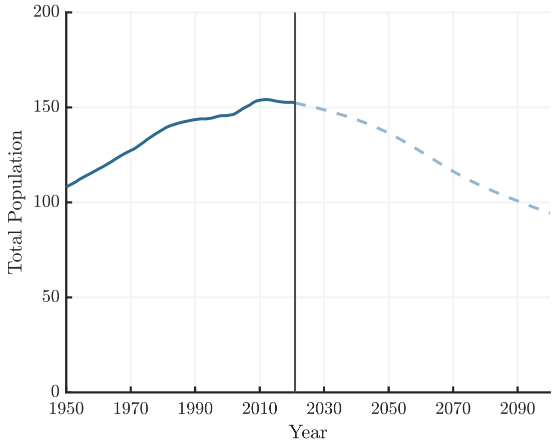
All said, Spain's demographic situation, while a bit more extreme than that of other similar countries, is still comparable in broad terms: TFR below replacement

[34]This Spanish baby-boom occurred later than that of other high-income countries, like the United States, where the baby-boom took place right after WWII and lasted until the mid-1960s.

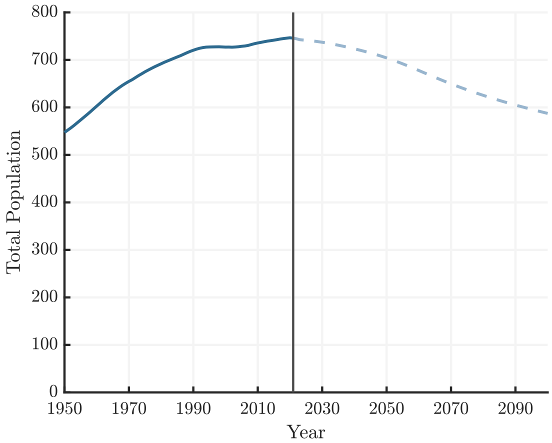
Figure A2.2 Total Population, Selected Countries (1950-2022)



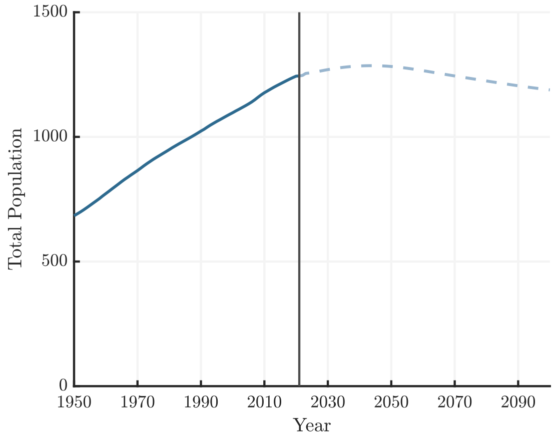
(a) Spain



(b) Southern Europe



(c) Europe



(d) High Income

Source: United Nations, Department of Economic and Social Affairs, Population Division (2022).
Note: Projections account for fertility, mortality, and migration, and we use the medium scenario. The total population is measured in thousands.

in all likelihood will lead to population aging and decline.

A2.2 Calibration Appendix

Labor Income Process

The parameters of the labor income process together with the persistence of the income shock and the variance of the residual of it were estimated from the Spanish Continuous Working Life Sample (MCVL). In particular, we regress gross log monthly income by gender on experience and experience square. We also control by age, education, year, province, occupation, sector, tenure, a dummy for part-time jobs, and the interaction between age and part-time jobs. After this, we estimate the residuals of both regressions assuming that they follow an AR(1) process as it is described in the model.

In this appendix, we summarize the main variables and sample restrictions that we made for the income process estimation. We use STATA codes from De La Roca and Puga (2017). [35]

1. **Main variables:**

    - **Gross monthly income:** refers to a very approximate measure of all labor income received by a person except pensions, prizes from games like the National Lottery, and non-levied income. Therefore it does not include unemployment benefits. It is extracted from the tax codes because they are uncensored. Earnings are expressed in real terms using the consumer price index of 2009.
    - **Education.** It is divided into three educational levels: less than secondary education, secondary education, and university education.

2. **Sample restrictions:**

    - We restricted the sample for the years 1998-2017. The main reason for this is that the specification over the type of contract (open-ended vs fixed-term contracts) is available after this year.
    - We keep the individuals who entered the labor market after 1998.
    - We dropped:
        (a) Individuals older than 55 years old. This eliminates those who receive an early retirement pension.
        (b) Unemployed individuals
        (c) Immigrants.
        (d) Public sector employees.

[35] We are responsible for the possible computational mistakes.

Table A2.1 Labor Income Process by Gender and Contract

	Male	Female FT	Female OE
Experience	0.0457***	0.0475***	0.0409***
	(316.07)	(163.13)	(195.67)
Experience^2	-0.00140***	-0.00267***	-0.00134***
	(-184.87)	(-123.99)	(-127.33)
Age (years)	0.0225***	0.0538***	0.0166***
	(94.28)	(134.25)	(51.41)
Age^2	-0.000320***	-0.000763***	-0.000233***
	(-89.42)	(-123.33)	(-51.06)
Secondary education	0.0633***	0.0116***	0.0461***
	(181.66)	(17.27)	(94.46)
Tertiary education	0.135***	0.0498***	0.169***
	(255.74)	(63.61)	(285.64)
Part-time contract	-0.554***	-0.393***	-0.441***
	(-172.55)	(-139.73)	(-165.01)
Constant	7.093***	6.348***	7.154***
	(1770.06)	(883.34)	(1229.60)
Observations	8846471	2996927	5327129
Month FE	Yes	Yes	Yes
Province	Yes	Yes	Yes
Occupation	Yes	Yes	Yes
Sector	Yes	Yes	Yes
Tenure	Yes	Yes	Yes
Age×Part_time_contract	Yes	Yes	Yes
Fixed-term_contract	Yes	No	No

t statistics in parentheses

Source: MCVL 1980-2016, FT: Fixed-term contract, OE: Open-ended contract

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2.2 Persistance AR(1) Process by Gender

	Male	Female
resid_male_lag	0.900***	
	(2489.44)	
resid_female_lag		0.906***
		(2683.50)
Observations	17170526	17170526

t statistics in parentheses

Source: MCVL 1980-2016

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2.3 Residual Income Shock by Gender

	Mean	Variance
Residual Male	-.0065449	.0348526
Residual Female	.009062	.0344275
Observations	17170526	

Source: MCVL 1980-2016

Table A2.4 Probability of Transition to Permanent Contract $\pi_z(x)$ by Years of Experience x

x	$\pi_z(x)$
0	0.066
1	0.092
2	0.121
3	0.151
4	0.182
5	0.212
6	0.239
7	0.263
8	0.284
9	0.303
10	0.319
11	0.335
12	0.351
13	0.369
14	0.388
15	0.410
16	0.433
17	0.458
18	0.485
19	0.512

A2.3 Computation of the Model

Computation With Taste Shocks

All of the decisions that women make in the model are discrete in nature. Unfortunately, this type of model tends to generate a jerky aggregate response to parameter changes, since individuals tend to change their decisions all at once unless there are vast amounts of heterogeneity.

To facilitate the numerical solution of the model, we include a taste shock to women's utility in every period. This helps by smoothing out labor force participation and fertility decisions. The shocks can be interpreted as unobserved state variables that add noise to the women's decisions. Moreover, the calibration and results are robust to their inclusion. For an in-depth discussion of this computational method, see Iskhakov et al. (2017).

Thus, we assume that in every period women receive a vector of additive-separable taste shocks μ. In periods when they can still have children and need to choose on pregnancy $b \in \{0, 1\}$ in addition to labor force participation $h \in \{0, \frac{1}{4}, \frac{1}{2}\}$, they receive a vector of six shocks, one for every element in $\{0, \frac{1}{4}, \frac{1}{2}\} \times \{0, 1\}$. In periods when they cannot have any more children and need only to choose labor force participation, they receive a vector of three shocks, one for every element in $\{0, \frac{1}{4}, \frac{1}{2}\}$:

$$\mu = \begin{cases} \left(\mu_{0,0}, \mu_{\frac{1}{4},0}, \mu_{\frac{1}{2},0}, \mu_{0,1}, \mu_{\frac{1}{4},1}, \mu_{\frac{1}{2},1}\right) & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \left(\mu_0, \mu_{\frac{1}{4}}, \mu_{\frac{1}{2}}\right) & \text{otherwise} \end{cases}$$

All of these shocks are i.i.d, drawn from an Extreme Value Type I distribution with scale parameter σ_μ.

The modified value function in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ is:

$$W_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}, \mu; N^*\right) = \begin{cases} \max\{W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) + \sigma_\mu \mu_{h,b}\}_{h\in\{0,\frac{1}{4},\frac{1}{2}\}, b\in\{0,1\}} & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \max\{W_j^{h}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) + \mu_{h}\}_{h\in\{0,\frac{1}{4},\frac{1}{2}\}} & \text{otherwise,} \end{cases}$$

where $W_j^{h,b}$ and W_j^h represent the value, ex-taste shock, of choosing labor force participation h and pregnancy status b for a woman in period j, or just labor force participation h, in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$:

$$W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = u^{h,b}\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right]$$
$$W_j\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = u^{h}\left(c, l, \mathbf{n}; N^*\right) + \beta \mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right],$$

where $\mathbb{E}^{\sigma_\mu}$ denotes the expectations over future taste shocks and in both cases, choice and states variables need to be retrieved from the constraints and laws of motion:

$$\begin{aligned}
& l = 1 - h - \xi(\mathbf{n}) \\
& c = I - \lambda 2h\left(n_1 + n_2\right) \\
& I = y^m + 2hy^f\left(1 - \mathbb{1}_{\{h=\frac{1}{2}\}}\phi\right) - T\left(y^m, y^f, h\right) \\
& \ln\left(y^f\right) = \eta_0^f + \Delta\eta_0^f \mathbb{1}_{z=1} + \left(\eta_1^f + \Delta\eta_1^f \mathbb{1}_{\{z=1\}}\right) x + \left(\eta_2^f + \Delta\eta_2^f \mathbb{1}_{\{z=1\}}\right) x^2 + \epsilon^f \\
& \ln\left(y^m\right) = \eta_0^m + \eta_1^m (j-1) + \eta_2^m (j-1)^2 + \epsilon^m \\
& \epsilon^{f\prime} = \phi^f \epsilon^f + \nu^f \\
& \epsilon^{m\prime} = \phi^m \epsilon^m + \nu^m \\
& \mathbf{n}' = \Lambda_j(\mathbf{n}, b) \\
& x' = \Pi_x(x, h) \\
& z' = \Pi_z(x, h, z).
\end{aligned}$$

The main consequence of introducing the taste shocks is that the policy function becomes probabilistic. Given the distribution assumed for them, the probability that a woman chooses pregnancy decision b and labor force participation h in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ when $j < 15$ and $N(\mathbf{n}) < 3$ is the logit probability:

$$P_j\left(h, b \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \frac{\exp\left(\frac{W_j^{h,b}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}}\sum_{k\in\{0,1\}}\exp\left(\frac{W_j^{i,k}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}.$$

Otherwise, the probability that a woman chooses labor force participation h in states $\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)$ is the logit probability:

$$P_j\left(h \mid \epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right) = \frac{\exp\left(\frac{W_j^{h}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}{\sum_{i\in\{0,1,2,3\}}\exp\left(\frac{W_j^{i}\left(\epsilon^f, \epsilon^m, z, x, \mathbf{n}; N^*\right)}{\sigma_\mu}\right)}.$$

One additional benefit of using Extreme Value Type I shocks is that the expected value function is given by the tractable log-sum formula from (McFadden, 1973):

$$\mathbb{E}^{\sigma_\mu}\left[W_{j+1}\left(\epsilon^{f\prime}, \epsilon^{m\prime}, z', x', \mathbf{n}', \mu; N^*\right)\right] =$$
$$\begin{cases} \sigma_\mu \log\left(\sum_{i\in\{0,1,2,3\}}\sum_{k\in\{0,1\}} \exp\left(\frac{W_j^{i,k}\left(\epsilon^f,\epsilon^m,z,x,\mathbf{n};N^*\right)}{\sigma_\mu}\right)\right) & \text{if } j < 15 \text{ and } N(\mathbf{n}) < 3 \\ \sigma_\mu \log\left(\sum_{i\in\{0,1,2,3\}} \exp\left(\frac{W_j^{i}\left(\epsilon^f,\epsilon^m,z,x,\mathbf{n};N^*\right)}{\sigma_\mu}\right)\right) & \text{otherwise.} \end{cases}$$

Using backward induction starting in period J, one can easily retrieve the expected value functions and the probabilistic policy functions.

A2.4 Stochastic Structure for Children's Aging

In this appendix we show how to retrieve the stochastic structure that governs the transition probabilities for the vector state of the number of children at different ages takes, which we denote by $\mathbf{n}' = \Lambda_j(\mathbf{n}, b)$.

We denote by λ_1 and λ_2 the probabilities that an individual baby becomes a school-age child, and that an individual school-age child becomes a teen in a given period, respectively. Moreover, we assume that the aging event is independent across children. Notice that $b\alpha_j$ is the probability that there is a newborn in the next period. Denote by $P_i(x \mid n_i)$ the probability that x children in stage $i \in \{1, 2\}$ (babies, school-age) move on to the next stage the next period (school-age, teenager), conditional on there being n_i children in that stage in the current period. Table A2.5 shows these, for babies and school-age children.

Table A2.5 Probabilities of aging by number of children, $P_i(x \mid n_i)$

n_i	Number of children aging			
	0	1	2	3
0	1	0	0	0
1	$1-\lambda_i$	λ_i	0	0
2	$(1-\lambda_i)^2$	$\lambda_i(1-\lambda_i)$	λ_i^2	0
3	λ_i^3	$\lambda_i(1-\lambda_i)^2$	$\lambda_i^2(1-\lambda_i)$	λ_i^3

To compute the whole set of probabilities of transition from state $\mathbf{n} = [n_0, n_1, n_2, n_3]$, we follow the algorithm:

for $x_1 \in \{0, 1, 2, 3\}$ **do**
  for $x_2 \in \{0, 1, 2, 3\}$ **do**
    if $n_0 = 1$ **then**
      $\mathbf{n}' = [0, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $(1 - b\alpha_j)P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$ or
      $\mathbf{n}' = [1, n_1 - x + 1, n_2 + x - y, n_3 + y]$ w.p. $b\alpha_j P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$;
    else
      $\mathbf{n}' = [0, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $(1 - b\alpha_j)P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$ or
      $\mathbf{n}' = [1, n_1 - x, n_2 + x - y, n_3 + y]$ w.p. $b\alpha_j P_1(x_1 \mid n_1)P_2(x_2 \mid n_2)$;
    end
  end
end

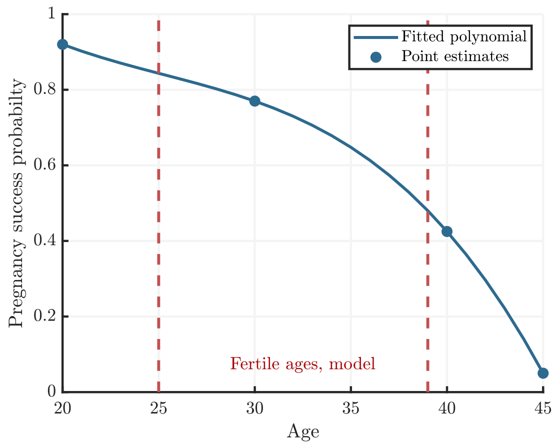
A2.5 Pregnancy Success Probabilities

We follow Sommer (2016) in estimating the probability of pregnancy success by age α_j. We use the following point estimates of natural infertility from Trussell and Wilson (1985):

$$\text{Infertility probability} = \begin{cases} 0.080 & \text{at age 20} \\ 0.230 & \text{at age 30} \\ 0.575 & \text{at age 40} \\ 0.950 & \text{at age 45,} \end{cases}$$

and then fit a polynomial through (the inverse of) them, as shown in Figure A2.3. Notice that we only have women having children in the model between ages 25 and 39.

Figure A2.3 Pregnancy Success Probability Conditional on Age

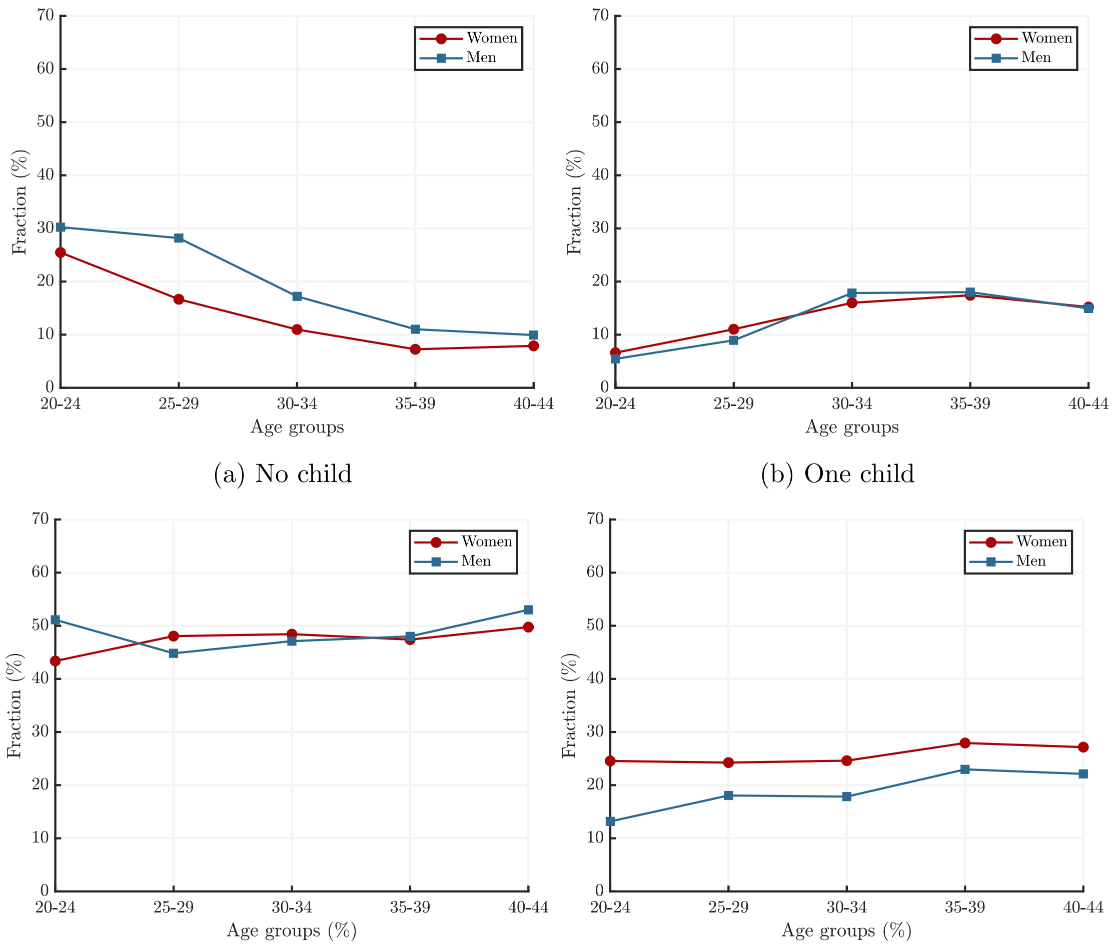


Source: Author's work, point estimates by Trussell and Wilson (1985).

A2.6 Men's Fertility Preferences

While the first wave of the Spanish Fertility Survey in 1999 sampled only women, the second one collected responses on men's fertility preferences. A disadvantage is that the samples are separate, i.e., the survey collected basic data on the spouses of the sampled individuals, but it did not collect data on their preferences. Nevertheless, it stands to reason that one of the characteristics that people sort on in marriage markets is desired overall number of desired children. Figure A2.4 shows fertility preferences by gender and age in 2018. The most significant discrepancies are the fractions of people wanting no children (higher for men) and those wanting three or more (higher than for women). However, in aggregate terms, the disagreement is minor and decreases with age.

FIGURE A2.4 Desired Fertility by Gender and Age in Spain



(a) No child (b) One child

(c) Two children (d) Three or more children

Source: *Encuesta de Fecundidad 2018, INE.*

Chapter 3

Gender Gaps in the Labor Market and Social Security Finances

3.1 Introduction

Pay-as-you-go pension systems have remained a key focus for policymakers for decades. Increasing longevity and decreasing fertility rates lead to demographic aging in societies, which have an impact on their financial sustainability. The old-age dependency ratio, the ratio of pensioners over contributors, is projected to increase from 30% in 2015 to 58% in 2075 for the EU28 (OECD, 2017). This will place additional burdens on the working-age population to finance pay-as-you-go pensions. As a result, researchers have studied the harmful implications of rapid population aging on pension finances and proposed pension reforms to alleviate the fiscal burden.[1]

Two other structural changes have also influenced the pension system concurrently with the aging process: the increase in the educational attainment of individuals and the rise in female employment rates. While some studies have explicitly analyzed the effect of education on the pension system (Conesa et al., 2020; Díaz-Saavedra, 2022), it has not yet been determined whether the increase in female employment rates helped to alleviate the financial burden generated by the aging population. This paper aims to answer the following question: Did women help alleviate the financial burden caused by the demographic transition? If so, by how much and for how long?

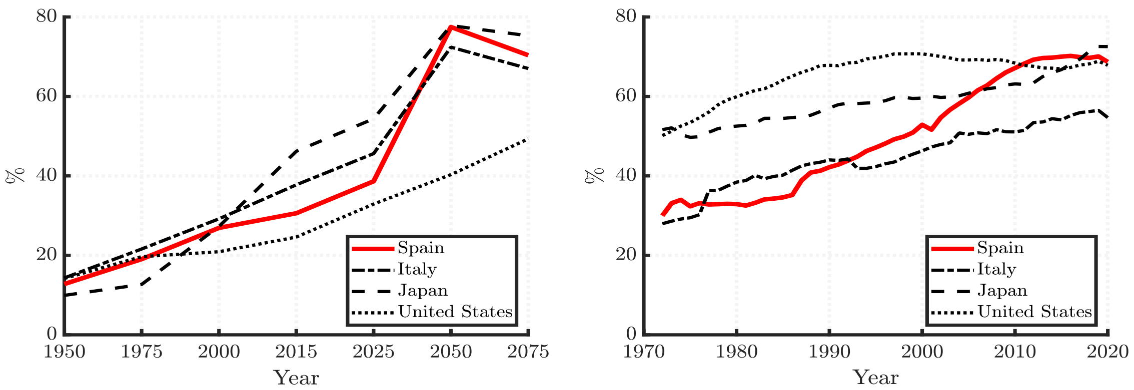
The Spanish economy is a suitable case study for three key reasons. First, Spain has emerged as one of the foremost advanced economies regarding the celerity and enormity of demographic aging.[2] The Old-age dependency ratio is projected

[1]For instance: De Nardi et al. (1999), Kotlikoff et al. (2001), Galasso and Profeta (2004), Attanasio et al. (2007), Kitao (2015) and Nishiyama (2015), among many others.

[2]A comparative analysis of Spain with other countries concerning the proportion of tertiary education, fertility rates, and life expectancy at age 65 by gender can be found in Figure A3.1

to escalate from 30.6% in 2015 to 70.4% in 2075, see Figure 3.1a. Second, the expenditure on public pensions represents a big share of the Spanish government finances (10.9% in 2019), and its sustainability is in trouble. The debt of Social Security reached a historical maximum of 6% of the Gross Domestic Product (GDP, henceforth) in 2020, and the previous pension reforms (2011 and 2013) do not seem to solve this fiscal problem completely.[3] Third, there has been a notable and swift surge in the participation of women in the labor force. While the female labor force participation in Spain stood at a mere 30% in 1975, it had risen steeply to 70% in 2019, surpassing leading countries such as the United States, as illustrated in Figure 3.1b.

FIGURE 3.1 Old Age Dependency Ratio and Female Labor Force Participation Trends in Selected Countries



(a) Old-Age Dependency Ratio (b) Female Labor Force Participation

Source: OECD statistics

Building on the seminal work of Auerbach and Kotlikoff (1987), I develop a deterministic, overlapping generations small open economy populated by married households and a government. Households are heterogenous in several dimensions, such as the age, the education of each spouse, the number of children, their productivity, and pension rights. When the household is born, the wife draws a participation cost to join the labor market and makes a career choice. This decision differentiates households into one and two-earner households, which I denote as "traditional" and "modern" couples, respectively. After that, both spouses solve a joint maximization problem where they decide their lifetime profile of consumption and each spouse's leisure. I assume the household has access to a perfect credit market where the household can borrow and save at a given interest rate. Furthermore, the model economy incorporates the main components of the Spanish Social Security regarding retirement and widow pensions. I assume that the budget of the pay-as-you-go

included in Appendix A3.1.

[3]See Díaz-Gimenez and Díaz-Saavedra (2017) and Conde-Ruiz and González (2015) for a discussion of these two reforms.

Social Security system balances at every period, and I compute the equilibrium path of the payroll tax rate.

I calibrate this model to an initial steady state in 1975 to reproduce the Spanish economy's basic demographic and economic properties. In 1976, unexpected changes in several demographic parameters trigger the transitional dynamics in labor supply, consumption, pensions, and social security contributions. In particular, I assume that starting in 1976, newborn generations have a higher life expectancy, higher educational attainment, fewer children, and a lower cost of labor market participation than the cohorts born before 1976. The participation cost of the new cohorts of women is calibrated so that the employment rate in 2019 is 20.6%, 37.4%, and 66.9% for low, medium, and highly-educated married women, respectively. After 2019, I use Spanish projections of demographic variables to solve the transitional dynamics towards a new steady state which is reached around the year 2100. At any period of this transition, the social security system's budget is balanced, which implies a large increase in the payroll tax, which reaches a maximum in the year 2050.

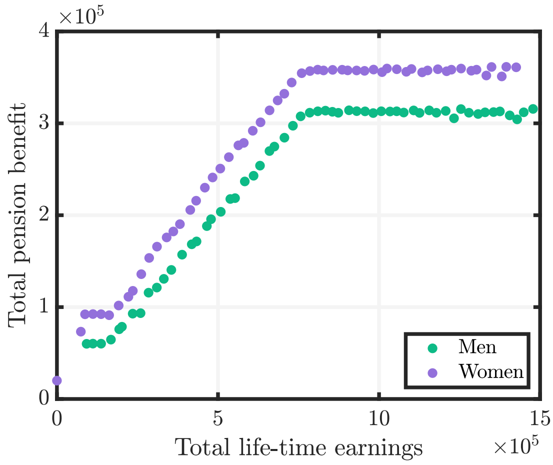
The main finding of this paper is that the Social Security system benefits from a “women bonus” which evaporates in the long run. To measure how much women contributed to the Social Security system, I created a novel indicator called the Female Gender Imbalances Indicator that quantifies the share of male pensions financed by women. In particular, women finance 10% of men's pensions in 2019 and continue to do so until 2050. After this year, men begin to finance women's pensions, implying that the women's increase in the Social Security system is only a temporary solution.

To understand the mechanism behind this result, in Figure 3.2, I plot the total lifetime earnings (x-axis) and total pension benefits (y-axis). The figure reveals two conclusions: firstly, benefits display a concave pattern, suggesting that including individuals with below-average incomes, specifically women, worsens the long-term situation. Secondly, due to widow pensions and longer lifespans, women receive higher benefits than men, further exacerbating the long-term challenges faced by the Social Security system.

Given the projections I use, I also find that the model converges to a final steady state where persistent labor market gaps exist, such as participation, work hours, earnings, and average pensions. Motivated by this and the fact that women's elasticity to labor supply is higher than men's, I quantify the effects of introducing Gender-Based Taxation (GBT) in Spain. In particular, I assume that after 2019 women are permanently taxed at a lower rate than men. This policy generates a considerable drop in those gaps and generates welfare gains for newborn cohorts.

The introduction of Gender-Based Taxation is a controversial topic due to various reasons. Firstly, imposing a tax system that disproportionately taxes women at

Figure 3.2 Total Lifetime Earnings and Total Pension Benefits in the Baseline Model



Note: Each dot represents the average pension benefit given the lifetime earnings for all individuals in the Baseline Model.

a lower rate than men could raise constitutional concerns in many countries. Secondly, while married women tend to have a more elastic labor supply than men, the labor supply elasticity of single individuals is more similar to that of men. Lastly, the elasticity of individuals' labor supply is an endogenous object. Differences in labor supply elasticities between genders "do not only depend on innate characteristics or preferences but may emerge endogenously from the internal organization of the family" Alesina et al. (2011). This makes the design of an endogenous GBT system even more challenging. Given these complexities and potential constitutional concerns, the paper avoids discussing the implementation of GBT as a policy option. Instead, it focuses on presenting theoretical results related to GBT and suggests that exploring alternative reforms that could achieve similar benefits may be a fruitful avenue for future research.

The paper lies at the intersection of three lines of research. The first concerns a long tradition in macroeconomics research quantifying the macroeconomic and redistributive effects of Social Security. This literature uses life-cycle models pioneered by Auerbach and Kotlikoff (1987). In particular, I contribute to the literature that analyzes pay-as-you-go pension systems. İmrohoroglu et al. (1995) studied the optimal replacement rate in the United States. De Nardi et al. (1999) quantify the impact of the demographic trends in the U.S. on the sustainability of the Social Security system. Fuster et al. (2007) evaluates the elimination of the Social Security System with altruistic individuals. In Spain, Díaz-Giménez and Díaz-Saavedra (2009) studied the effect of delaying the statutory legal retirement age. However, I depart from the existing literature by considering a two-earner household environment and modeling the female labor supply. To the best of my knowledge, Sánchez Martín and Sánchez-Marcos (2010) are the first and only authors who introduced this new household's environment in the pension literature. However, modeling females' labor supply in the extensive and intensive margins is widespread in other set-ups. For instance, Guner et al. (2012a) studies the effect of two proposed tax reforms in the U.S. in a life cycle model with married and single households. Kaygusuz (2015) and Nishiyama (2019) assess the implications of removing the Old Age and Survivor's program of the U.S. Social Security system on the labor supply of married women.

The second relevant strand of literature pertains to recent research concerning the growth in female labor force participation observed over previous decades. This branch of literature can be divided into two areas. The first area focuses on documenting this phenomenon for the US (Goldin (2006)) and Spain (Guner et al. (2014)). The second area proposes several explanations behind the changes in women's labor market outcomes. On the supply side, various studies suggest that technological advancements in households that reduce the cost of home-produced goods (Greenwood et al. (2005)), the availability of contraceptive pills (Goldin and Katz (2002)), advancements in medical technologies related to motherhood such as infant formula (Albanesi and Olivetti (2016)), a reduction in the cost of childcare and the gender wage gap (Attanasio et al. (2008)), have contributed to this change. Other potential explanations include the importance of cultural beliefs (Fernández and Fogli (2009)), a decrease in discrimination against women (Jones et al. (2015)), or an increase in aggregate productivity in professional occupations (Gayle and Golan (2012)). In light of this, the present paper extends the literature cited above to a new section that explores the consequences of this labor market trend on Social Security finances.

The third strand of the literature is the evaluation of tax reforms with a heterogeneous agent dynamic framework, for instance, Ventura (1999), Díaz Giménez and Pijoan-Mas (2006), Conesa and Krueger (2006), Kaygusuz (2015), and Guner et al. (2012a), among many others. Guner et al. (2012b) quantify the effect of taxing women at a lower rate than men in the U.S. They find that setting a proportional tax rate on married females increases output and female labor force participation and generates welfare gains. Different to Guner et al. (2012b), I quantify the effects of introducing gender-based taxation into the Spanish Social Security system.

The remainder of the paper is organized as follows. Section 2 describes the institutional framework. Section 3 describes the socioeconomic trends in Spain. Section 4 describes my quantitative model. Section 5 specifies the calibration methodology in the initial steady state and presents the model results. Section 6 describes the transition, calibration and results. Section 7 computes the role of women in Social Security funding. Section 8 shows the results and the welfare analysis of introducing Gender Based Taxation. Finally, Section 9 concludes.

3.2 Institutional framework

The pension program designed for the elderly population in Spain is classified as a pay-as-you-go defined benefit system. The fundamental components of the prevailing pension structure were initially established in 1967 and further improved throughout the 1970s. Over the past 50 years, the system has undergone six significant reforms in 1985, 1997, 2002, 2007, 2011, and 2013. The contributory pension scheme comprises four categories: old age, disability, survivor, and orphanhood. Furthermore, three primary regimes for pensions exist the general regime, special schemes, and government employee schemes.

This paper concentrates on the general regime (RGSS), which accounted for 71% of the retirement pensions in Spain in January 2019. In what follows, I briefly outline the principal pension regulations for old-age pensions under the RGSS and their main changes over time.[4]

Financing. The pension system is financed by contributions, which are fixed percentages of covered earnings (total earnings minus bonus payments) between a floor and a ceiling. In 2019, workers and employers must contribute to the Social Security system 4.7% and 23.6% of taxable earnings to finance pensions.

Eligibility. To be eligible for a retirement pension, the worker must have contributed to Social Security at least N_c years before retirement. Before 1985, $N_c = 8$, but after 1985 and until today, the minimum contributed years are 15. Before the 2011 reform, the legal retirement age was 65, but it has now been raised to 67. Early and late retirements are allowed.[5][6] However, late retirement is rewarded.

Pension formula. If eligible, the pension is calculated by multiplying the replacement rate by the regulatory base ("regulatory base"). The regulatory base is a weighted average of monthly covered earnings over a reference period comprising the last N_b years before retirement. Contributory years have changed over time. Before 1985, only two years were considered, then eight in 1985, fifteen in 1997, and 25 in 2011. The replacement rate decreases with the years the worker retires before the legal retirement age and increases with the years the worker retires after the legal retirement age. As the employee reaches retirement age, the replacement rate

[4] I show the main changes in Table A3.1 in the Appendix A3.1; however, see Garcia-Mandicó and Jiménez-Martín (2020) for a detailed description.

[5] Before 1985, early retirement was only available to workers who made their first contribution before 1967. In 2002, early retirement at age 61 was extended to the rest of the population. The 2011 reform further restricted early retirement.

[6] Except for workers in dangerous professions who find it challenging to maintain their working activities, early retirement is penalized.

increases by the number of years he or she has contributed to Social Security. A worker must have worked N_f years to qualify for the full pension, which was always 35 years until the 2011 pension reform, which increased it to 37 years.

Maximum and minimum pensions. Since 2002, pensions have been subject to an annual legislated ceiling roughly equal to the covered earnings ceiling.

The existing literature largely agrees that implementing these pension reforms alone cannot effectively address Spain's Social Security financing issue.[7] The Social Security debt will likely constitute a substantial portion of the Spanish GDP over time. In the subsequent section, I will delineate the key socioeconomic trends substantiating this claim.

3.3 Socioeconomic Trends in Spain

This section provides an overview of the evolution of employment rates, life expectancy for men and women, fertility rates, and education transitions in Spain between 1975 and 2019. Further, it discusses the main projections in these areas and their implications for the sustainability of the Spanish pension system.

3.3.1 Labor Market Trends

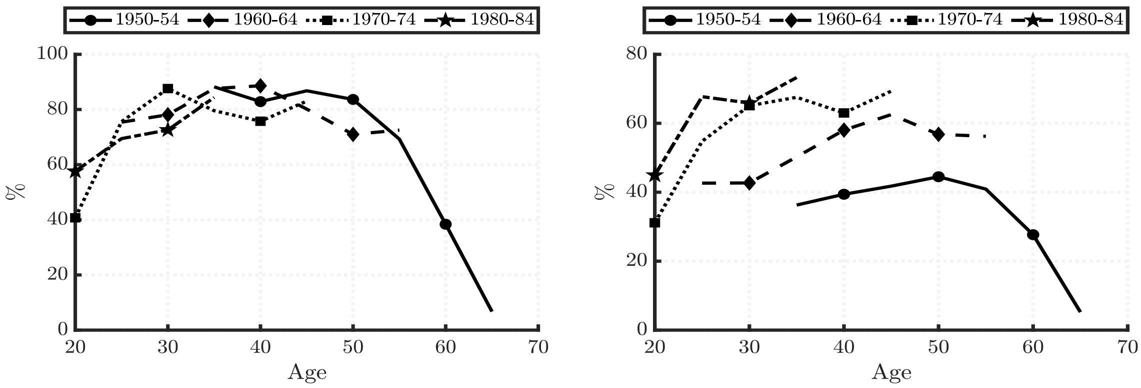
During the authoritarian regime (1939-1975) in Spain, the Catholic Church endorsed the male-breadwinner model, which limited women's roles to childcare and household tasks. Married women required their husbands' permission for many legal and financial matters. However, the end of the dictatorship in 1975 marked a turning point for women's rights. The advent of democracy, entry into the European Union in 1986, and social and policy changes, including anti-discrimination laws and free public schooling for children aged three and above in the 1990s, helped transform the cultural norms surrounding women's roles.

Since the 1970s, the male-breadwinner model has been gradually supplanted by the dual-breadwinner model, where women have paid jobs, contribute to the household income, and thereby alter the traditional gender roles within the household. Figure 3.3 illustrates the employment rates by gender for four cohorts: 1950-54, 1960-64, 1970-74, and 1980-84. The left panel of Figure 3.3a indicates that men's employment rates remained relatively constant across these cohorts. However, there was a significant increase in women's employment rates, as displayed in the right panel of Figure 3.3b, with a considerable rise across all age groups. For example,

[7] See Appendix A3.2 for a literature review.

at age 35, 36% of women born in the 1950s were employed, compared to 73% for women born in the 1980s.

Figure 3.3 Employment Rate by Gender and Cohort

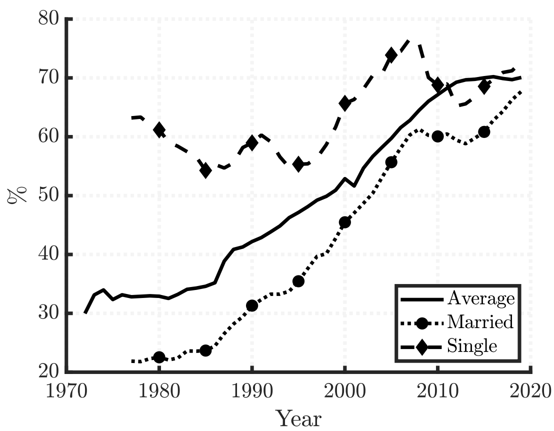


(a) Men (b) Women

Source: Spanish Statistical Office (INE)

Figure 3.4 illustrates the evolution of the female employment rate over time compared to the rates observed for single and married women. The figure shows that the female employment rate increased from 30% in 1975 to 70% in 2019, mainly due to the higher employment rate among married women. Specifically, single women experienced a 9.5 percentage point increase, while married women experienced a 45.8 percentage point increase. These trends suggest a significant shift in women's labor market behavior, which Guner et al. (2014) have analyzed in more detail.

Figure 3.4 Female Employment Rate by Marital Status and the Economy's Average



Source: Spanish Statistical Office (INE).

3.3.2 Demographic Trends

During the last few decades, many developed countries have undergone a demographic transition characterized by declining fertility rates and increasing life expectancy. In this new phase, countries face challenges related to an aging population, where a smaller working-age population must support a larger elderly population. Spain is one of the leading countries experiencing this demographic shift, with a low fertility rate and a significant increase in life expectancy.

Life expectancy. In recent decades, men and women have experienced a rapid increase in life expectancy at 65. In the early 1970s, 65 years old men and women had 14 and 17 years of life expectancy, respectively. In 2019, life expectancy at age 65 increased to 19 and 22 years for men and women, respectively. According to the Spanish Statistical Office, by 2050, men and women are expected to have a life expectancy at age 65 of 21.9 and 25.5 years, respectively. This demographic shift alone will lead to an extension of pension payments and a subsequent increase in pension expenditure.

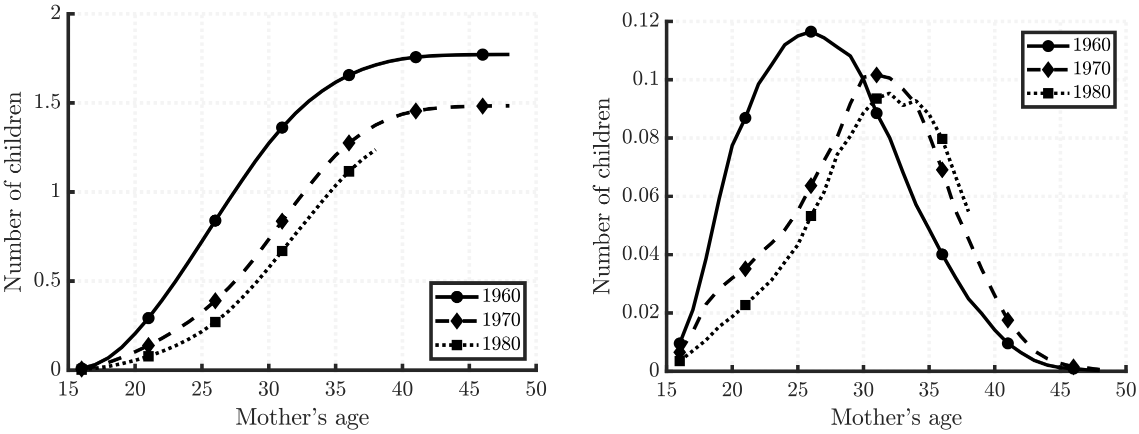
Fertility rates. It is well known that fertility rates have plummeted, and women have been delaying fertility. To better give a magnitude of those trends, I make a cohort analysis of the fertility rates using the birth rates between 1975 and 2018 provided by the Spanish Statistical Office. I focus on three cohorts of women: born in the 1960s, 1970s, and 1980s. The 1960s and the 1980s are the oldest and the youngest generations, for which I observe their completed fertility rates.

Figure 3.5 presents the cumulative cohort fertility rate in Figure 3.5a and the age-specific fertility rates in Figure 3.5b. The results reveal that recent cohorts exhibit a declining trend in the number of children. Specifically, women in the 1960s cohort had an average of 1.8 children at the age of 48, whereas those born in 1980 had 1.24 children at the age of 38. Additionally, the postponement in the onset of childbearing is approximately five years.[8]

I merged the aforementioned data with Eurostat's projections to calculate the cohort fertility rate projections. The evolution of these projections is presented in Figure 3.6. On average, the fertility rate declines from 1.8 children per household for women born in 1959 to 1.46 for those born in 2019. Although there is a slight increase in total fertility rates after the 1993 birth cohort, projections indicate that the levels will still be considerably below those of women in the 1960s cohort.

[8]Read Ahn and Sánchez-Marcos (2020) for a complete analysis of the fertility rates by cohorts in Spain.

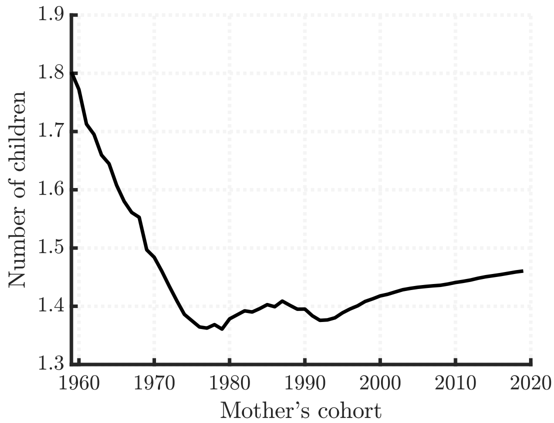
Figure 3.5 Cumulative Cohort and Age-Specific Fertility Rates



(a) Cumulative Cohort Fertility Rate (b) Age-Specific Fertility Rate

Source: Birth records 1975-2018 from INE.
Note: The fertility rate is defined as the per-period total births by 1000 mothers in fertile ages (16-48)

Figure 3.6 Cohort Fertility Rate

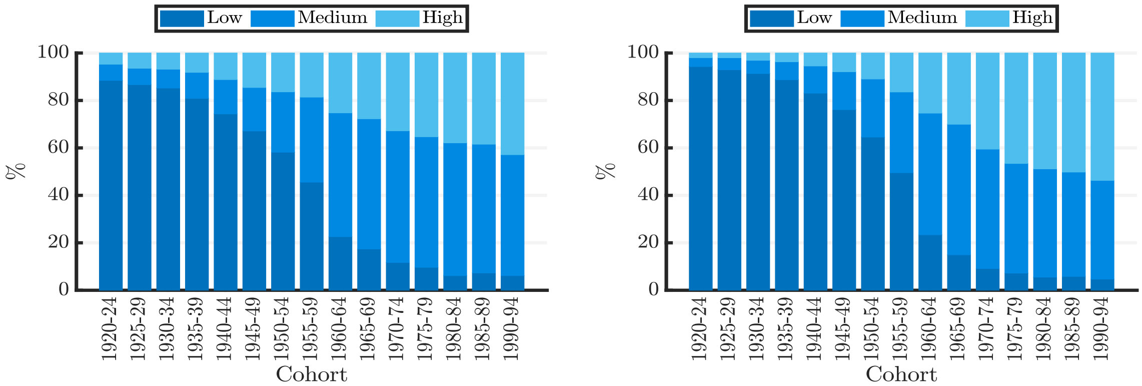


Source: Birth records 1975-2018 from INE and projections 2019-2067 from Eurostat.
Note: The fertility rate is defined as the per-period total births by 1000 mothers in fertile ages (16-48).

3.3.3 Educational Trends

Spain has witnessed rapid and extensive growth in the educational qualifications of its population. Figure 3.7 shows the educational attainment levels of men and women across cohorts characterized as low, medium, and high education. The cohort-wise analysis of educational attainment indicates that until the '60s cohort, the typical characteristic of the Spanish population was a low level of education, with approximately 74% and 80% of men and women born between 1920-60 having low educational attainment. However, there has been a shift in this trend, with merely 6.2% and 4.7% of men and women born between 1990-94 falling into the low educational attainment category. Furthermore, women born in the first half of the 1990s reveal higher educational attainment than men, with 54% being highly educated as opposed to 44% of men.

Figure 3.7 Educational Attainment by Gender and Cohort



(a) Men (b) Women

Source: INE.

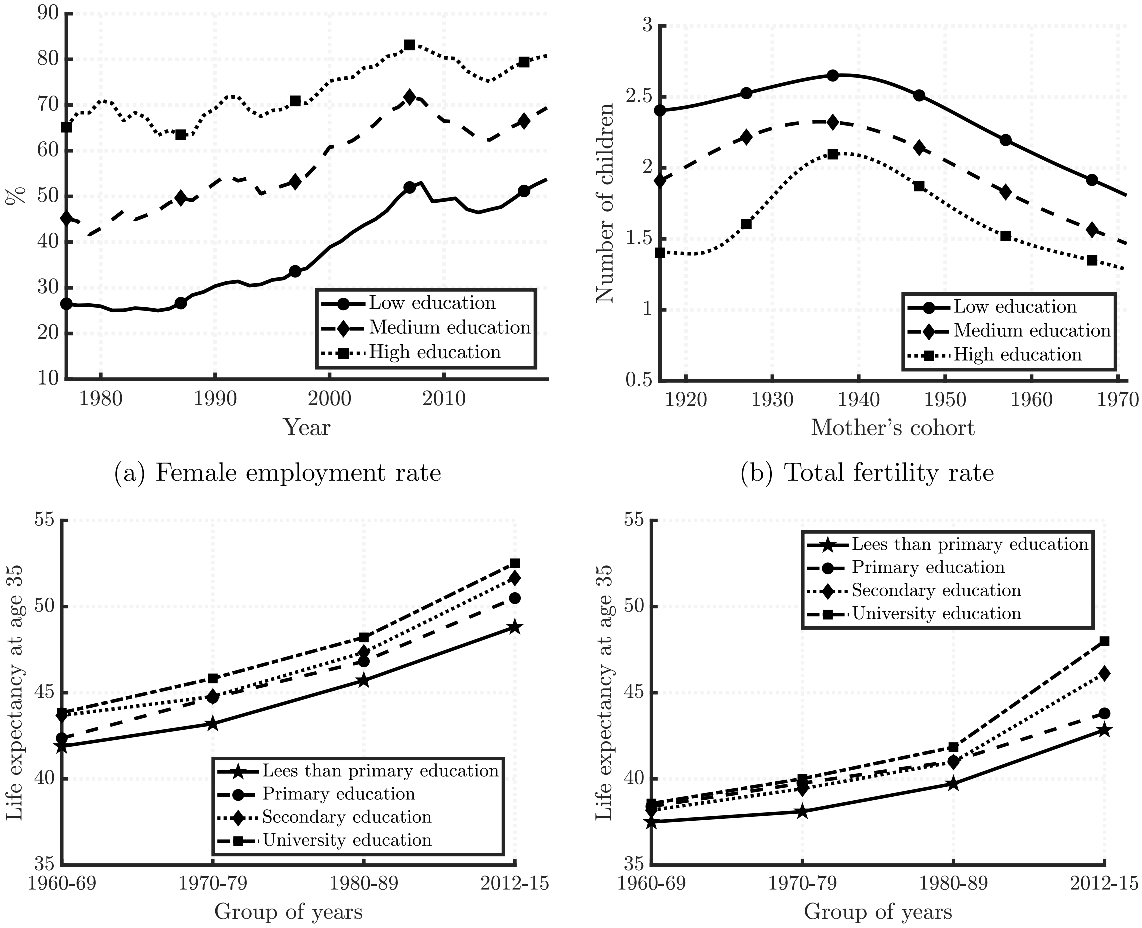
Note: A low-education individual is someone with, at most, a primary education level, whereas a medium-educated individual has completed secondary education, including upper secondary education and vocational training. High-educated individuals have a tertiary education.

Changes in the composition of women's education impact the labor market, fertility rates, and life expectancy. Figure 3.8 shows the evolution of the female employment rate over time, the total fertility rates by education and cohort, and the life expectancies over time and gender. Figure 3.8a shows that the female employment rate increases with the level of education. The rise in the proportion of women attaining tertiary education, in combination with this relationship, has a positive impact on the average female employment rate. Notably, the employment rate of low-educated women experienced the most substantial increase between 1977 and 2019, narrowing the employment gap by education. Figure 3.8b shows that the total fertility rate has declined across all educational attainments. Thus, differences in the fertility rate across education are persistent. Since education negatively correlates with the number of children, increasing women's education also impacts the low aggregate fertility rates.

Regarding education, Permanyer et al. (2018) analyzed the relationship between educational attainment and life expectancy in Spain from 1960 to 2015. Figures 3.8c and 3.8d shows the estimated life expectancy for women and men by education, respectively. Individuals with higher levels of education generally experience longer life expectancies than those with lower levels of education. This gap has continued to widen over time, as evidenced by the fact that although life expectancy at age 35 has increased for all education groups, it has increased more rapidly for the highly educated. Specifically, from 1960 to 2050, the life expectancy at age 35 increased by

6.9 years for women with low education and 8.7 years for those with high education. For men, the corresponding increases were 5.3 years and 9.4 years, respectively. Higher education correlates with longer life expectancy, so the compositional change favoring more educated individuals will also impact future life expectancy.

FIGURE 3.8 Female Employment, Total Fertility Rates, and Life Expectancy at Age 35 by Education



(a) Female employment rate

(b) Total fertility rate

(c) Life expectancy for women

(d) Life expectancy for men

Source: INE For the female employment rate, Zeman et al. (2014) for the total fertility rate and Permanyer et al. (2018) for the life expectancy.

3.3.4 Implications of the Trends in the Social Security Finances

The demographic changes we have observed in Spain, namely the increase in life expectancy and the decline in fertility rates, indicate that the Spanish society is aging. This shift in the population structure will have a detrimental impact on pension sustainability. The primary reason for this is the rise in the proportion of retired cohorts relative to working-age cohorts, as evidenced by the increase in the old-age dependency ratio.

The dependency ratio–the ratio of the population older than 65 to the population of working age (16-64)– increased from 16.62% to 29.9% between 1975 and 2019. Demographic projections indicate that the dependency ratio will be 57% in 2050. This demographic shift will further exacerbate the challenges of financing pensions, as the pool of contributors will become smaller, and the burden on those remaining will increase.

At the same time, the rise in the population's educational attainment also impacts Social Security funding, apart from its effects on employment, fertility, and life expectancies discussed in the previous section. Workers with a college degree usually have higher productivity compared to non-college-educated ones. This will positively increase the tax revenues of Social Security but also will increase Social Security expenditure because they will claim higher pensions compared to their non-college-educated.

In a nutshell, the male-breadwinner model has been gradually replaced by the dual-breadwinner model, with a significant increase in women's employment rates across all age groups. The female employment rate increased from 30% in 1975 to 70% in 2019, mainly due to the higher employment rate among married women. Fertility rates have declined, and women have been delaying fertility. Life expectancy has significantly increased for both men and women, which will lead to an increase in pension expenditure. These demographic shifts present challenges related to an aging population, where a smaller working-age population must support a larger elderly population.

3.4 The Model Economy

I study an overlapping generations model economy with a continuum of heterogeneous households and a government. The unit of analysis is the household, which consists of a husband and wife of the same age who never divorce. I assume Spain is a small open economy, and the model is deterministic. Time is discrete and runs forever. In every period t, a new household generation, s, is born. Throughout this section, I present the model economy in a steady state, where variables do not depend on the period.

3.4.1 The Households

Households are heterogeneous concerning age j, the educational attainment of the husband and the wife z, x, the number of children $\phi(x)$, the initial labor market productivity of each spouse ϵ^i; the household type m; the labor market status of each spouse s; the private assets accumulation b and pension rights of each spouse e. I denote $i \in \{w, h\}$ to refer to the wife and the husband, respectively.

Age. Individuals enter the economy at age 20 ($j = 1$) as workers and retire at the exogenous age of 65 ($j = J_R$). Life is deterministic, and it lasts J_h periods for the husband and J_w periods for the wife, where $J_h < J_w$. Women in the model live longer than men, so I only model widow pensions for women.

Education. At the beginning of life, each spouse is endowed with an exogenous type that refers to educational attainment and remains constant over the life cycle. For husbands, the type is z, where $z \in Z$ and $Z \subset R_{++}$ is a finite set. For wives, the education type is given by x, where $x \in X$ and $X \subset R_{++}$ is a finite set. I consider three levels of education for both genders: low, medium, and high. I denote the distribution of education type of couples by $M(z, x)$, where $\sum_{z \in Z} \sum_{x \in X} M(z, x) = 1$.

Children. Each household has an exogenous number of children determined by the wife's educational attainment, $\phi(x)$. I assume that all the children arrive deterministically to the household. Children aged 0-2 (babies) impose both time (ψ) and monetary costs (q). While the time cost is levied on all mothers in the economy, the monetary cost only applies to working mothers and is proportional to their labor supply. I denote by $\bar{\phi}_j(x)$ the number of babies in the household at age j.

Labor market productivity. The labor market productivity of each age,j, spouse, i, by education type has two components. First, when individuals enter the economy, they draw an initial productivity level from a log-normal distribution denoted by ϵ^i, which depends on the individual's educational attainment. It satisfies the following conditions: $\epsilon^i \in E$, where $E \subset R_{++}$ is a finite set. I denote the time-invariant distribution of the initial productivity of the spouses by $S\left(\epsilon^h(z), \epsilon^w(x)\right)$. Second, education and gender determine the individual's earnings profile over the life cycle. $\varpi_j^h(z)$ and $\varpi_j^w(x)$ denote the wage at age j of a husband with education z and wife with education x, respectively.

The gross labor income of a worker i at age j supplying l_j^i units of labor and, given the economy wage rate ω, is given by:

$$\begin{aligned} y_j^h &= \varpi_j^h(z)\epsilon^h(z)\omega l_j^h \\ y_j^w &= \varpi_j^w(x)\epsilon^w(x)\omega l_j^w. \end{aligned} \tag{3.1}$$

Household type. At the beginning of life, women decide whether to stay home and constitute a traditional couple or to join the labor market and form a modern couple, $m \in \{0,1\}$. To make this decision, in line with Guner et al. (2012a), I assume that in the first period, wives draw a participation cost, κ, from a cumulative distribution function that depends on the wife's education type $\Omega(\kappa|x)$. This cost aims to account for the additional constraints women face when deciding to enter the labor market, and I assume that it is constant over the life cycle. $\Omega(\kappa|x)$ denotes the probability of the utility cost being κ and $\sum_{\kappa \in K} \Omega(\kappa|x) = 1$. Women are less likely to work when these costs are sufficiently high, while women are more likely to work when these costs are low.

Labor market status. In this economy, there are three labor market statuses: employed, non-active, and retired, $s\{e,n,r\}$. Employed individuals are in the labor market and receive a salary that depends on the efficiency units and hours worked. In this model, I assume all husbands are employed during the working state, $j \in [1, J_R]$, while only women in modern households are employed. Non-active individuals refer to women without a job and who are not looking for a job, i.e., only women in traditional households are in this status. All individuals, after $j = J_R + 1$, become retirees. This decision is exogenous in the model. Whether they are entitled to a retirement pension or a widow pension depends on the pension rights described below.

Private assets. All households enter the economy with no assets and die without assets. There is a perfect credit market where households can borrow and save at the interest rate of r. Therefore, during the life cycle, there is heterogeneity in the asset holdings, $b \in B$.

Pension rights. Individuals differ in their pension rights. For retirement pensions, the Spanish pension system determines that eligibility depends on the amount of contributed years to the labor market, N_c. Moreover, for widow pensions, the law states that all wives who become widows receive a widow pension which is independent of their retirement pension rights. I assume that agents are myopic about pension rights and the pension formula.

Preferences. Following Erosa et al. (2022), preferences are represented by a per-period utility function that depends on the husband and wife's consumption and leisure. Notice that I assume the same Pareto weight of husbands and wives in the household utility. The per period utility function of a household at age j is:

$$
\begin{aligned}
U(c_j, l_j^h, l_j^w) = & 2\ln(c_j) + \nu \frac{(1-l_j^h)^{1-\gamma}}{1-\gamma} + \nu \frac{(1-l_j^w - \bar{\phi}_j \psi)^{1-\gamma}}{1-\gamma} \\
& - \kappa m,
\end{aligned} \tag{3.2}
$$

where c_j is the public good household's consumption and l_j^i is the labor supply of each spouse i. In modern households, $m = 1$, and females incur a utility cost κ. If a female has a baby, she also incurs a time cost proportional to the number of babies in the household $\bar{\phi}_j \psi$. Note that $\frac{1}{\gamma}$ is the intertemporal elasticity of leisure, and ν is the weight on work's disutility, common across genders.

3.4.2 The Firm

In this model economy, I assume that there is a representative firm. It produces output with standard constant returns to scale labor-augmenting production function:

$$Y_t = K_t^{\theta}(A_t L_t)^{1-\theta} \tag{3.3}$$

where (A_t, K_t, L_t) represents the productivity, the aggregate capital, and aggregate labor, respectively. Capital depreciates at a constant rate, δ. The firm rents capital in the international capital market at an exogenous rate $r + \delta$ and hires domestic labor supply at a wage per efficiency unit of ω. The law of motion of the labor augmenting productivity factor A_t is $A_{t+1} = (1 + \widetilde{g})A_t$, where the technological process is exogenous.

3.4.3 The Government

The government in this economy runs a pay-as-you-go pension system. It taxes labor earnings with a payroll tax from the working-age cohorts.[9] It uses these revenues to provide pensions to the retired cohorts. The Social Security budget constraint in period t is:

$$P_t = T_t, \tag{3.4}$$

where P_t denotes aggregate pensions in period t and T_t, is the aggregate revenues collected by the payroll tax in period t. I assume the government uses the payroll tax, τ_t, to balance the Social Security budget for all periods.[10] In Appendix A3.3, I carefully describe how I aggregate pensions and contributions.

Retirement pensions. To qualify for a retirement pension in Spain, an individual must have contributed to the Social Security system for at least N_c years. When a retiree has the right to receive a pension, the amount depends primarily on the regulatory base and the replacement rate.

[9] I assume that the payroll tax is not capped.
[10] I do not model the reserve fund introduced in Spain in 2000.

The regulatory base refers to an individual's average earnings over the past N_b periods before retirement. The regulatory base is multiplied by the replacement rate, which increases with the working years. Additionally, retirement pensions are bounded below and above to capture the minimum and maximum pensions. Worker i's retirement pension is calculated according to the following formula:

$$p^i = \alpha^i \quad \frac{\sum_{j=J_R-N_b}^{J_R-1} y_j^i l_j^i}{N_b}, \tag{3.5}$$

where N_b is the number of years considered for computing the regulatory base (average lifetime earnings), and α^i is the replacement rate. The pension system replacement rate is a step function that increases with the total individual's number of worked years at the retirement age, x^i. In particular, I use the following formula:

$$\alpha^i = \begin{cases} 0 & \text{if } x^i < a_1 \\ \alpha_0^p + \triangle\alpha_1^p(x^i - a_1) & \text{if } x^i \in [a_1, a_2] \\ \alpha_2^p + \triangle\alpha_3^p(x^i - a_2) & \text{if } x^i \in [a_2, a_3] \\ 100\% & \text{if } x^i > a_3. \end{cases} \tag{3.6}$$

Widow pensions. Following the death of her husband, the wife receives a widow pension proportional to his retirement pension. Any retired female with a retirement pension lower than the maximum pension is eligible for this widow pension. The widow's pension is determined by a fixed proportion χ of her husband's pension:

$$pd = \chi p^h. \tag{3.7}$$

There are specific bounds for both retirement and widow pensions in the model. For example, $\bar{P}$ and $\bar{Pd}$ represent the maximum quantities for retirement and widow pensions, while $\underline{P}$ and $\underline{Pd}$ represent their minimum quantities.

3.4.4 The household's problem

In this section, I describe the household problem of a household when it is born, given the household type m. I will omit the dependence of variables on the husband and wife's education and time. The household maximization in present value, assuming that all variables are in efficiency units of labor, is described by:

$$\max_{\{c_j\}_{j=1}^{J_w},\{l_j^h,l_j^w\}_{j=1}^{J_R}} \sum_{j=1}^{J_R} \beta^{j-1}\Big[2\ln(c_j) + \nu\frac{(1-l_j^h)^{1-\gamma}}{1-\gamma} + \nu\frac{(1-ml_j^w-\bar{\phi}_j\psi)^{1-\gamma}}{1-\gamma}\Big] - \kappa m + \sum_{j=J_R+1}^{J_h}\beta^{j-1}\Big[2\ln(c_j)\Big] + \sum_{j=J_h+1}^{J_w}\beta^{j-1}\Big[\ln(c_j)\Big]$$

subject to

$$\begin{aligned}
&\sum_{j=1}^{J_w}\frac{c_j(1+g)^{j-1}}{(1+r)^{j-1}} = m\sum_{j=1}^{J_R}\frac{\left((1-\tau_j)l_j^w y_j^w - q\bar{\phi}_j l_j^w\right)(1+g)^{j-1}}{(1+r)^{j-1}} \\
&+\sum_{j=1}^{J_R}\frac{\left((1-\tau_j)l_j^h y_j^h\right)(1+g)^{j-1}}{(1+r)^{j-1}} + \sum_{j=J_R+1}^{J_h}\frac{p_j^h(1+g)^{j-1}}{(1+r)^{j-1}} \\
&+m\sum_{j=J_R+1}^{J_w}\frac{p_j^w(1+g)^{j-1}}{(1+r)^{j-1}} + \sum_{j=J_h+1}^{J_w}\frac{pd_j(1+g)^{j-1}}{(1+r)^{j-1}} \\
&c_j \geq 0 \\
&l_j^h \in [0,1] \\
&l_j^w \in [0,1] \quad \text{if } m=1; \quad l_j^w = 0 \quad \text{if } m=0
\end{aligned} \tag{3.8}$$

3.5 Calibration Details: Initial Steady State

This section describes the initial steady-state calibration. The model economy's functional forms and parameters are intended to reproduce the Spanish economy in 1975 and the Spanish pension system. The calibration consists of a two-step estimation strategy. In the first step, I assign values to parameters that can be identified outside the model. Then, in the second step, I calibrate the remaining parameters in the second step to target some moments of the Spanish economy. In Appendix A3.1, Table A3.2 summarizes the central parameter values for the initial steady state.

3.5.1 Parameter Values Set Exogenously

Timing. The length of the model period corresponds to one calendar year. The agent's endowment of discretionary time is 5200 hours per year.[11] Individuals' disposable time is split into labor, leisure, and raising babies in the case of women.

Demographics. Lifetime is certain and common to all generations alive in the initial steady state. Individuals enter the economy at age 20 and die deterministically. The lifetime of individuals in the initial steady state is taken from the life

[11] I consider 16 hours per day, 6 days per week, and 52 weeks per year of discretionary time.

expectancy at age 65 in 1975 provided by the Spanish Statistical Office (INE, the acronym in Spanish). In the initial steady state, women live until age 82, $J_w = 63$, and men until age 79, $J_h = 60$. Retirement is deterministic and happens at age 65, $J_R = 46$.

The initial share of age groups in the population, $N_{t,j}$, corresponds with the Spanish population pyramid by age from INE in 1975. This implies an old-age dependency ratio of 17.2% while the data counterpart is 16.82%.[12]

Education. Individuals' education levels are classified as low, medium, and high. In the initial steady state, each cohort has specific educational attainment. To compute the initial distribution of households by the educational attainment of the spouses, I follow the methodology in Esteve and Cortina (2006) using the Spanish Census of 2001 and 2011. I carefully explain the methodology in Appendix A3.4, and in Table A3.4, I show the calibration results. In the initial steady state, 59% of households are composed of two spouses with low education, 19% of medium education, and 4% of high education.

Initial productivity. At the beginning of the life cycle, individuals draw an initial productivity from a log-normal distribution that is gender and education specific. I estimate the mean and the variance of the distributions using the "Muestra Continua de Vidas Laborales" (MCVL) and the Industry and Services Wage Survey. Since in the 70s, there was a lot of selection of women for those in the labor market, I assume that the variable of female distribution is the same as men's, and for the mean, I calibrate it to match the exogenous gender wage gap specified below. Table 3.1 shows the mean and variance of the log-normal distribution by education. See Appendix A3.4 for a complete description of the sample restrictions and the specific methodology.

Table 3.1 Mean and Variance of the Log-Normal Distribution by Education

	Mean, μ^h	**Variance, σ**
Led	9.42	0.327
Med	9.56	0.396
Hed	9.76	0.450

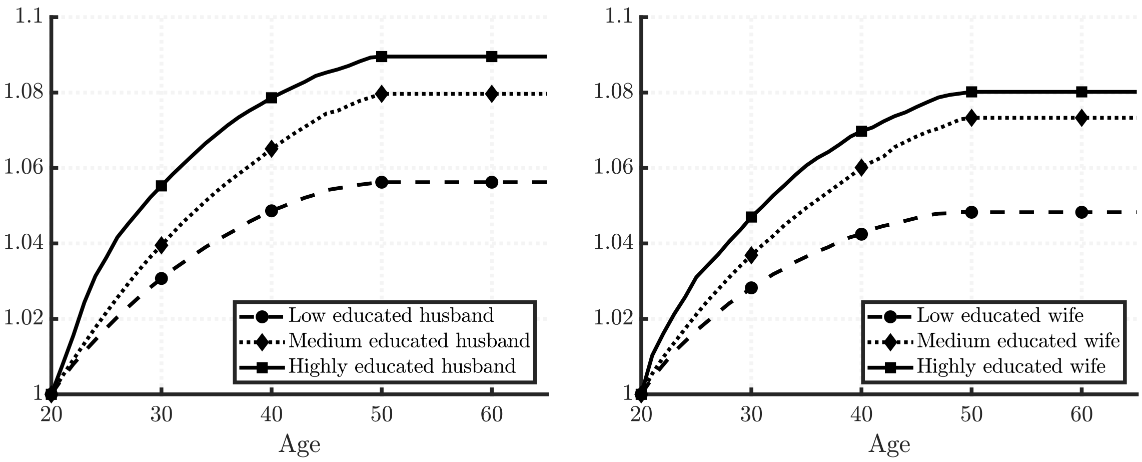
Slope of earnings. I estimate the earnings profile over the life cycle by gender and educational attainment using the MCVL. In particular, the slope of the earnings profile comes from the predicted values of the following estimated model:

[12] The model and the data are not perfectly matched because people enter the economy at the age of 20, whereas the data include individuals from the age of 16, and some people's lifetimes are longer in reality than those in the model.

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \tag{3.9}$$

where Y_i are the log yearly earnings of an individual i. X_i is a control vector that includes: age, age squared, a dummy variable for low and medium education, an interaction term between education and age, and a dummy variable for gender and occupation. The slope of life-cycle earnings is normalized to one for all individual types. See Figure 3.9 for a gender and education comparison.

Figure 3.9 Predicted Profile of Earnings by Education and Gender in the Initial Steady State.



(a) Husbands (b) Wifes

Source: MCVL, using the 1980-2001 sample.

Return of assets. For the asset's return, r, which is the interest rate of the world's credit market, I use the real one-year Treasury Bill rate between 1960 and 1975. In particular, $r = 0.036$, which is in line with the computations of Licandro et al. (1998) for Spain (3.75%) and Erosa et al. (2016a) for U.S. (4%).[13]

The productivity growth. I set the labor augmenting productivity growth equal to the average growth rate of real Gross Domestic Product per capita between 1913 and 1975 estimated by De la Escosura (2017). This implies a productivity growth of $g = 1.77\%$.

Number of children. The number of children the household has is mother's education-specific. To compute these fertility rates, I use the cumulative fertility

[13]There is no consensus in the literature about which is the best proxy for the world interest rate. "The 3-month U.S. T-Bill rate, the rate of return on the S&P 500, the LIBOR rate, and a weighted average of several countries' T-Bill rates, have been employed as nominal interest rates" Blankenau et al. (2001).

rates of the cohorts 1910-1955 by education of the mother from the Zeman et al. (2014) database. The number of children attached to low, medium, and highly educated women in the initial steady state are 2.5, 2.2, and 1.8, respectively. See Appendix A3.4 for a complete explanation of the computation of these.

Arrival of children to the household. I compute the average age at first birth and the average spacing between children using the INE. In 1975, women had, on average, their first children at the age of 25, $J_c = 6$. Moreover, the second and the third arrive with a three and two years gap, i.e., $j = 9$ and $j = 11$, respectively.

The time cost of babies. Babies, aged 0-2, impose a time cost on mothers, independent of their labor market status. I estimate this cost using the INE's Spanish Time Use Survey (2001-2002). I set the time cost equal to the gap in the time spent with children between married mothers and fathers. This gap is 46 minutes per day, which in model terms implies that $\psi = 0.046$.

The monetary cost of babies. When the mother works, the household incurs in a monetary cost that accounts for the childcare cost. This cost is proportional to the mother's labor supply. To estimate the out-of-pocket childcare cost, I combined three different data sources: Immervoll and Barber (2006) analysis for 2002 with the OECD database, the Household Budget Survey provided by INE in 2006, and the survey of Household's Expenditure in Education provided by the INE in the course 2011-2012. These estimations were 6%, 7.4%, and 7.12% of full-time workers' average earnings, respectively. As there is no richer database, I set the per-child care cost faced by full-time working women, q is the 7% of average earnings in the economy.

The pension formula. Using the General Regime System of 2002, I determine that the retirement age equals 65 years old, $J_R = 46$. To qualify for a retirement pension, a worker must have worked full-time for at least 15 years, $N_c = 15$. For a complete distinction between full-time and part-time employees, see Appendix A3.4. Regulatory bases consider gross labor earnings over the preceding 15 years to retirement, $N_b = 15$, while widows receive 52% of their husband's pension. In addition, I use the replacement rate parameters presented in Equation 3.6.

Retirement and widow's pensions are limited to guarantee a minimum income level at retirement and prevent getting too high pensions. These bounds are calculated as percentages of average earnings. I use the Social Security annexes from 1980-2014 and the average yearly earnings provided by De la Fuente (2017) to compute the average of these bounds. Table A3.2 in Appendix A3.1 summarizes all pension formula parameters.

Household preferences. The intertemporal elasticity of leisure $\frac{1}{\gamma}$ is set exogenously. I take the calibration in Erosa et al. (2022), so $\gamma = 4$. This parametrization implies that the Frisch elasticity of labor supply for a full-time worker ($l = 0.4$) is 0.375, which is between the estimates of the micro literature, see Erosa et al. (2016a).[14]

3.5.2 Parameter Values Set by Solving the Model

Six parameters remain undefined: the initial distribution of female wages, the taste of leisure, the three participation costs, and the discount factor. These parameters are chosen to minimize a loss function determined by deviations between the model implied and data moments. In what follows, I specify how I pin down each of the parameters, even though equilibrium outcomes determine them all. Table 3.2 shows the second-stage parameters and data moments.

Table 3.2 Second-Stage Parameters and Data Moments

	Parameter	Value	Targeted data moments
Mean female initial earnings	μ^w	$\mu^h - 0.27$	Gender wage gap full-time workers
Taste of leisure	ν	0.58	Average male working hours in 1987
Participation cost, led	$\varphi(1)$	9.9	Female employment rate led (25-45)
Participation cost, med	$\varphi(2)$	7.9	Female employment rate med (25-45)
Participation cost, hed	$\varphi(3)$	5	Female employment rate hed (25-45)
Discount factor	β	0.992	Wealth over income ratio

Note: led, med, and hed refer to low, medium, and highly educated women.

Initial productivity for females. For females, I assume the moments of the log-normal distribution of the initial productivity are the same as for males, but with a lower mean, μ^w. I calibrate this drop in the mean to match the gender wage gap for full time-workers to 17%, which is the estimation with the MCVL in 2001.[15]

Taste of leisure. I calibrate the taste of leisure, ν, to match the average hours worked by men during the working ages. The first available estimation from the INE is from 1987 and equals 39.1 hours per week. Since the total discretionary time in a year is 5200 hours (100 hours per week), the average fraction of time worked is 0.391. This calibration implies that the labor supply elasticity is 0.39 for males in the initial steady state.

[14] The Frisch elasticity of labor supply is $\frac{1}{\gamma}\frac{(1-n)}{n}$ where n is the average time worked.

[15] This estimation is in line with Guner et al. (2014) of 17.1% in 1994.

Participation cost. The calibration strategy of $\kappa(x)$ follows Kaygusuz (2010), Kaygusuz (2015) and Guner et al. (2012a). When the household enters the model economy, the wife draws a utility cost of participating in the labor market, $\kappa(x)$, from a gamma distribution that depends on her educational attainment. Once this cost is set, it is constant over the life cycle. The scale parameter η is 5 for all education groups. The shape parameter $\varphi(x)$ is calibrated to match the women's employment rate by educational attainment in the initial steady state. As estimated by Sánchez-Marcos (2003), the female employment rate for married women aged 25-45 in 1980 for low, medium, and highly educated women was 20.7%, 37.3%, and 66.5%.

The discount factor. Following Kaplan and Violante (2010), the discount factor is calibrated to match the median household wealth over income ratio in 2002. I take the definition and data from the Encuesta Financiera de las Familias (EFF) reported by the Bank of Spain in 2002. Net wealth is defined as the total value of assets (both real and financial) minus the number of debts. Income is defined as the sum of the property income from the households' asset holdings and the gross labor and non-labor earnings received by all household members. The Spanish median wealth over income ratio was 4.32 in 2002. This estimation is much higher than the one reported in Kaplan and Violante (2010) for the U.S. economy. The explanation relies on the high wealth of Spain compared to the U.S., driven mainly by housing. As there is no available data for the previous years, I assume that the evolution of the wealth-to-income ratio behaves as the evolution of personal wealth over national income in Blanco et al. (2021). In particular, from 1975 and 2002 it grows from 400% to 500%. Therefore, the implied wealth-to-income ratio is 3.46 for the year 1975.

3.5.3 Calibration Results for the Initial Steady State

In this section, I show how well the model matches the target moments explained before. I also study the model's ability to reproduce some untargeted moments.

Targeted Moments

As shown in Table 3.3 the model perfectly matches the moments targeted by the initial steady-state calibration in 1975. At the aggregate level, the model fits the wealth-to-income ratio, the average hours worked by males, and the gender wage gap in the Spanish economy in 1975 satisfactorily. The model also almost perfectly matches the employment rates of married women by education, which are targeted by modifying the shape of the gamma distributions for the participation costs women face if they want to join the labor force.

Table 3.3 Model Fit: Targeted Moments in the Initial Steady State

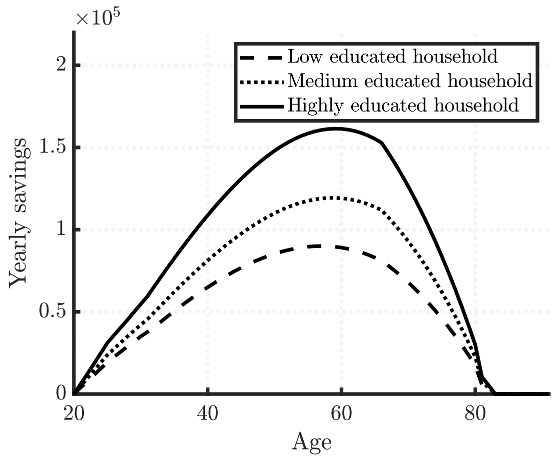
		Model year	Model	Data
	Led	1975	20.6%	20.7%
Married female employment rate (25-45)	Med	1975	37.4%	37.3%
	Hed	1975	66.9%	66.5%
Male average weekly working hours		1975	39.1	39.5
Wealth over income ratio		1975	3.3	3.5
Gender wage gap of full time workers		1975	17.6%	17.0%

Note: led, med, and hed refer to low, medium, and highly educated women.

Implications for Untargeted Statistics

In the initial steady state, the model generates a plausible hump shape profile of assets over the life cycle. Figure 3.10 shows the evolution of assets over the life cycle for three households composed of a wife and a husband with the same education. Savings are a tool for smoothing consumption since the model has no aggregate uncertainty. In traditional households, women are not entitled to a retirement pension, so this mechanism is essential.

Figure 3.10 Life-cycle Profile of Assets in the Initial Steady State



In the calibration, I target the employment rate of married women, but I do not target the number of hours worked. Table 3.4 shows that the model reasonably replicates the share of married women working more than 35 hours per week for medium and highly-educated women in 1980. However, the model implies that 23.09% of low-educated women work more than 35 hours per week, while 14% do in the data. The lack of career interruptions, apart from childcare, might explain why their hours worked are higher among low-educated women than in the data.

Table 3.4 Share of Women Working More than 35 Hours per Week in the Initial Steady State.

	Year	**Model**	**Data**
Low educated	1980	23.09%	14%
Medium educated	1980	25.46%	23.10%
Highly educated	1980	31.30%	32.30%

Source: Sánchez-Marcos (2003)

3.6 Aging Transition with PAYG Pension System

This section shows the model's transition between steady states under a PAYG pension system which assumes a budget balance each period. In Section 3.6.1, I describe the demographic, educational, and participation costs of women and the fiscal scenarios I use in my simulation. Section 3.6.2 discusses the calibration results for the targeted and untargeted moments. Section 3.6.3 provides an overview of the evolution of the Social Security contributions rate, household composition, and gender gaps in employment, hours worked, earnings, and pensions between steady states.

3.6.1 Methodology: Projections and Calibration

The benchmark and the counterfactual economies have the same initial conditions described in Section 3.5. Below I describe my assumptions during the transition. I divide the description into exogenous and endogenous changes.

Exogenous changes

Age distribution. I use the age distribution of households and its projections from the INE between 1976 and 2069. After 2069, I assume that it is constant. The age distribution implies that the old-age dependency ratio changes over time and grows from 17.78% in 1975 to 50% in 2069 (see Figure 3.11a).

Life expectancy. I assume that cohorts alive in the initial steady state make their decisions assuming that the life expectancy for men and women is 79 and 82 years, respectively. However, in 1976, I assume that the cohorts who are still in the working-age stage, receive a shock about their real life expectancy and reoptimize their decisions accordingly. Figure 3.11b shows the life expectancy I assume for cohorts alive in the initial steady state by their age at that time.[16] Newborn gener-

[16]For example, take the case of a man of age 20 in 1975. In 1975 he makes decisions according to the life expectancy of 79 and in 1976, he receives a shock stating that his new life expectancy

ations also face a cohort-specific life expectancy by gender, shown in Figure 3.11c. I assume that after 2014, it is constant. This implies that the life expectancy of men and women entering the economy after 2014 is 88 and 91, respectively. I select these life expectancies in Figures 3.11b and 3.11c to match the average life expectancy at age 65 for each period between 1976-2069. Since life is deterministic in the model, these parameters can be cleanly identified outside the model. See Figure A3.2 for the model fit.

Number of children and their arrival. Consistent with the data, I assume that the fertility rates by educational attainment of the mother drop for newborn cohorts. For cohorts born until the model period 1991 (the 1971's generation), I follow the same methodology as in the calibration of the initial steady state. For younger cohorts, I build the cohort fertility rate in the same fashion as Ahn and Sánchez-Marcos (2020), given the Birth Records for the 1972-2027 generations by the INE. Since this database does not provide fertility rates by education, I assume that the differences in the cohort fertility rates across educations are constant and equal to the 1971 cohort.[17] The number of children attached to low, medium, and highly educated women drop from 2.5,2.2 and 1.8 in 1975 to 1.8,1.5, and 1.3 in 2050, respectively. See Figure 3.11d for an evolution of the fertility rates by cohort. Also, the mother's age when giving birth to her first child increases consistently with the INE, assuming it remains constant after 2015. It increases from 25 in 1975 to 31 for cohorts arriving in the economy after 2015. See Figure 3.11e for the evolution of the mother's age at first birth.

Education. For cohorts born in the model years 1976-1998, I follow the same methodology and data source as in the initial steady-state calibration. For cohorts born between 1999-2015, I combine the data from the INE on the educational attainment of men and women with the Spanish Census. With the Spanish Census data, I construct a likelihood modifier that is informative about the assortative mating of couples. See Section A3.4 in the Appendix for a detailed description of the methodology and Table A3.3 for the evolution of the educational attainment by cohort. For cohorts born after 2015, I assume their educational composition is the same as in 2015.[18] The shares of couples with two spouses being drop-outs, high school graduates, and college graduates change from 59%,19%, and 4% in 1975 to 2%, 40%, and 29% in 2050.[19] Figure 3.12 shows the implied evolution of educational

is 85.

[17]Assuming the same differences in the fertility rates by education levels is not a stringent assumption. These differences are constant for the last ten available cohorts in the Zeman et al. (2014) database.

[18]Table A3.4 in the Appendix summarizes the implied share of household types through time.

[19]For educational attainment, my scenario aligns with those in the literature. For example, Díaz-Saavedra et al. (2023) assumes that men drop-outs, high school graduates, and college

attainment by gender.

Endogenous changes

Participation cost of women. I change the gamma distribution shape associated with the female participation cost between 1975 and 2019 to match the female employment rate by educational attainment of married women aged 25-54 in 2019. From 1976 to 2019, I assume the participation cost by education decreases linearly. Figure 3.11f shows the evolution of the gamma distribution of the participation cost by education.

Fiscal policy. I assume that the Social Security budget is balanced each period, and this determines a path of tax rates from 1976 on. I assume that individuals have perfect foresight about the path of taxes and the demographic changes since 1976.

3.6.2 Calibration Results

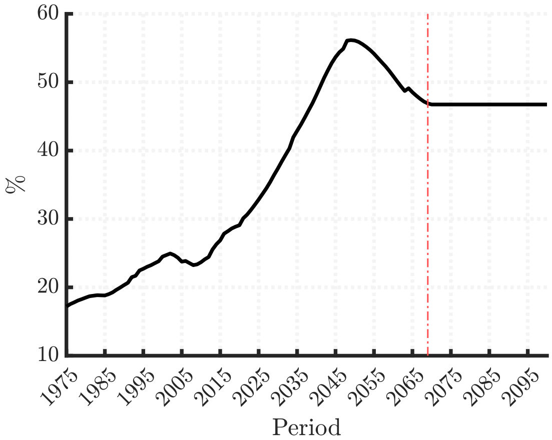
Targeted moments. Figure 3.13 shows the model fit of the employment rate by education of married women in 2019. This Figure shows that I match well the employment rates in 2019. Despite these being the only targeted moments, the linearly decreasing cost from the initial steady state to 2019 allows me to capture an evolution that is close to the data.

Implications for untargeted moments. Table 3.5 shows that the model successfully replicates selected labor supply and pensions statistics for 2019. It captures the effective hours worked by gender, the share of female pensioners, and the gender wage gap for full-time workers. The model does not replicate the correct size of the gender gap in retirement pensions and Social Security contributions. One should be cautious when comparing these gaps. The data moments represent the gender gap in contributions and pensions across all individuals, including singles, while the model only considers married individuals.

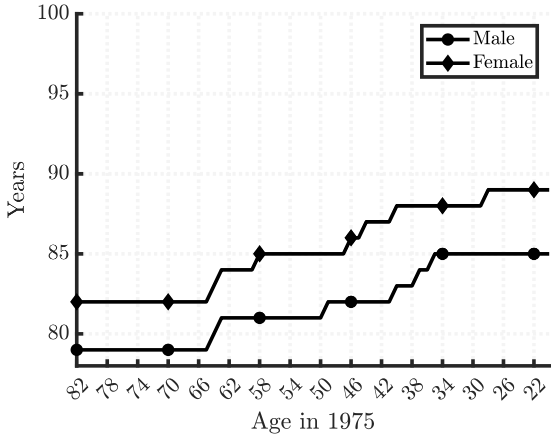
The gender gap in retirement pensions in the model is 6.2 percentage points lower than in the data. One possible explanation might be the fact that career interruptions are absent, except for children's arrival. In reality, prolonged absences from the labor market affect women's pensions. As a result, the model may generate a higher replacement rate than the data. The minimum pension may also be overestimated. This would imply that the model raises the pensions of low-income pensioners, who are most likely women.

graduates in 2050 will be 8.9%, 65.1%, and 26.0%, while I use 7%, 62%, and 32%, respectively.

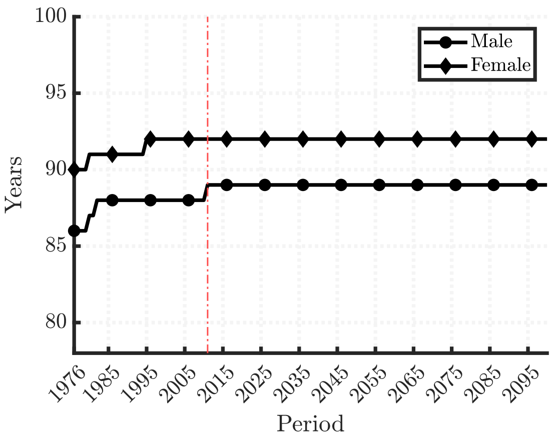
Figure 3.11 Scenarios Along the Transition



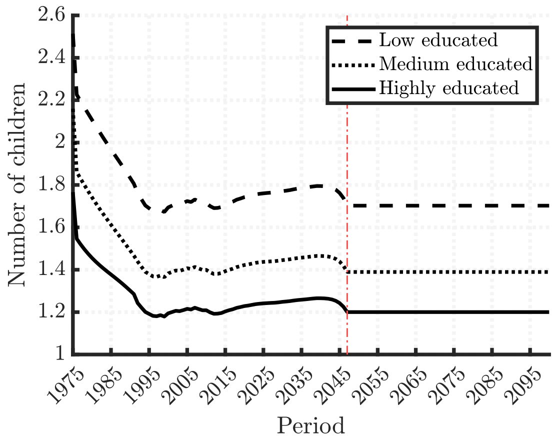
(a) Old-Age Dependency Ratio



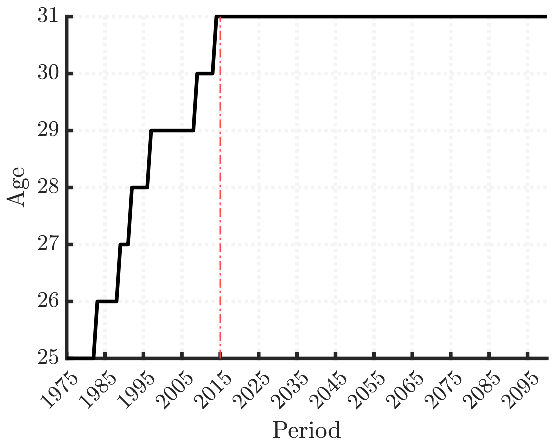
(b) Life Expectancy Cohorts Alive in 1975



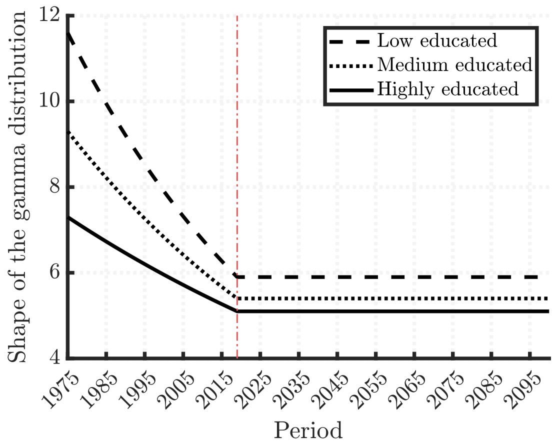
(c) Life Expectancy Newborn Cohorts



(d) Number of Children of Newborn Cohorts



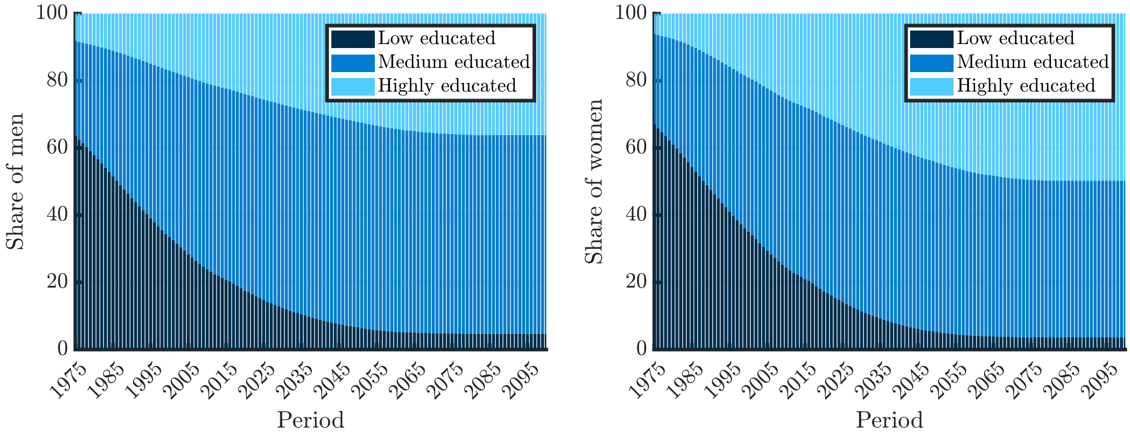
(e) Mother's Age at Birth Child Birth



(f) Shape of the Gamma Distribution

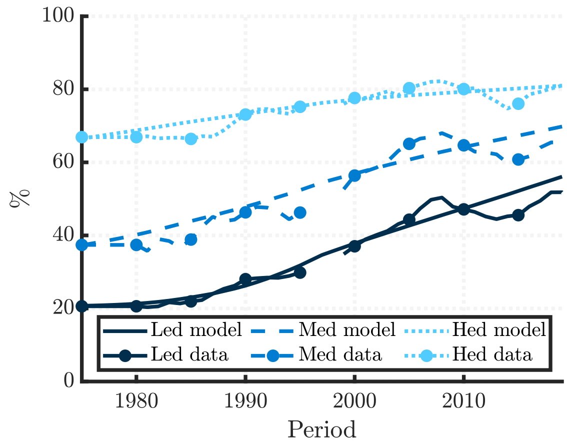
Note: The vertical line indicates the cohort from which the specific scenario is constant.

Figure 3.12 Evolution of Educational Attainment by Gender Along the Transition



(a) Share of Men by Education (b) Share of Women by Education

Figure 3.13 Model Fit: Female Employment Rate by Education



The model predicts that women contribute far less to Social Security than men. It overestimates this gap by 13.5 percentage points. The lack of single women may be the most plausible explanation. Women in this group work more and contribute more than married women. In 2019, their employment rate is about 70%. The lack of a minimum taxable income could also cause this mismatch. The only salaries contributing to Social Security are those above the minimum taxable income bracket. The model, however, taxes all income. Due to the positive exogenous gender wage gap, women are more likely to be contributing to Social Security in the model while they would not be eligible taxpayers.

3.6.3 Description of the Transition from 1975 to 2019 and Beyond

The economy is at an initial steady state in 1975. At this time, most households are traditional. In particular, 62% of them feature a husband that works and a wife

Table 3.5 Labor Market and Pension Outcomes in 2019

	Year	Model	Data
Female weekly hours worked	2019	31.45	30.4[a]
Male weekly hours worked	2019	36.89	36.2
Share of female pensioners over all pensioners	2019	36.71%	35.38%[b]
Gender gap in average retirement pension	2019	23.31%	29.52%[b]
Gender gap in average contributions to Social Security	2019	29.67%	16.2%[c]
Gender wage gap	2019 (2014)	12.87%	12.74%[d]

Note: [a] From the INE and excludes commuting time. [b] From the National Institute of Social Security (NISS), only considers retirement pensions to the General Regime (RGSS). [c] From the NISS and considers the difference between the average contribution base to the RGSS. [d] From Anghel et al. (2019).

who stays at home. The remaining 38% of households are modern, meaning women participate in the labor market. In the initial steady state, conditional on working, men work, on average, 29 hours more than women. This is due to several factors. First, women face a participation cost that prevents some women from participating in the labor market. Second, earnings are exogenous and are lower for women than for men. Third, having a child raises women's opportunity cost of working. Finally, if women work, they have to incur childcare costs.

Starting in 1976, the economy undergoes a transition that concludes in a new steady state in 2100. During the transition, several exogenous changes occur, reflecting what we observe in the data for the period between 1976 to 2019 and the projections I described in Section 3.6.1. First, both spouses' education levels increase. Second, fertility rates decrease for all women, regardless of their education. Third, women delay fertility. Fourth, men and women live longer. Finally, the age distribution shifts to older ages, resulting in an increase in the old-age dependency ratio.

Between 1976 and 2100, two endogenous changes accompany the exogenous changes above. First, the Social Security tax rate changes to balance the Social Security budget each period. Second, women's participation cost decreases linearly to match the female employment rate by education in 2019 and remains constant after that.

Figure 3.14 shows the transition path of the Social Security contributions and the old-age dependency ratio. The old-age dependency ratio rises steadily between 1976 to 2050, going from 17.8% to 59%. It then declines and stabilizes at 50% after 2069. As a result, the Social Security contribution rate increases between 1976 and 2050, going from 9.6% to 30.8%, and then lowers to 25.9% in 2100.[20]

[20]My results align with the pension literature. For example, Díaz-Saavedra et al. (2023) finds that the payroll tax rate would reach 51.1% in 2068 to fund the pension payments in Spain under

Figure 3.14 Social Security Contribution Tax Rate and the Old-Age Dependency Ratio in the Benchmark Economy

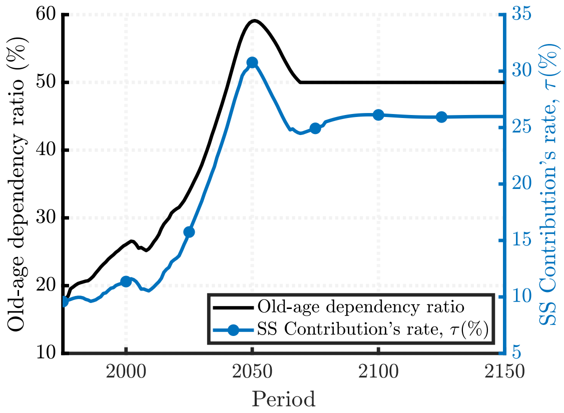


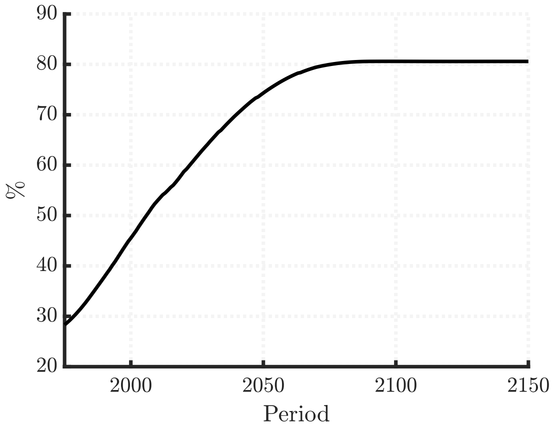
Figure 3.15 displays the transition paths of the share of modern households and hours worked by gender. Panel (a) shows that the share of modern households increases rapidly and reaches 80.6% in 2100. This increase is driven by several factors. First, the increase in the average level of education generates an incentive for women to participate in the labor market and form modern households. Second, the decrease in fertility rates implies lowers childcare costs and thus reduces the opportunity cost of creating a modern household. Third, the reduction in women's participation reduces the barriers women face when forming a modern household. Then, Panel (b) shows that, on average, women work 15.9 more hours in 2100 than in 1976. This is because women have fewer children, regardless of their education level. Furthermore, the composition of the female workforce changes towards highly educated women, who have a steeper life-cycle profile of earnings and higher productivity compared to their less educated counterparts. Finally, Panel (c) indicates that, on average, men reduce their weekly hours worked by 5 hours between steady states.[21] This reduction is driven by the increase in the Social Security contribution rate and the rise of modern households. In particular, compared to those in traditional households, husbands in modern ones work fewer hours. Additionally, during the transition, women in modern households work more, allowing husbands to work fewer hours.

Figure 3.16 displays the evolution of gender gaps during the transition. Panels (a)-(c) show that the gaps in employment rates, average working hours, and net earnings decrease steadily between 1976 and 2100. In particular, they reach mini-

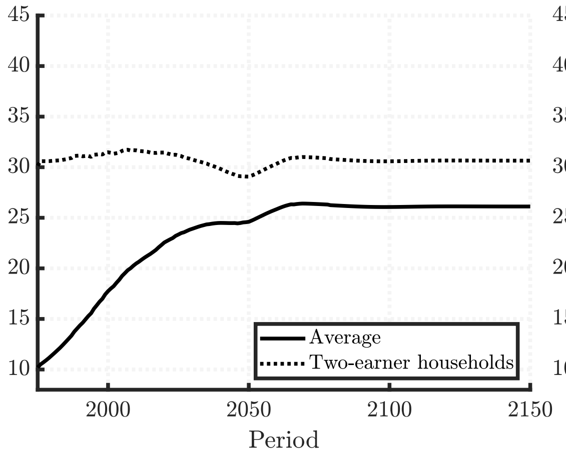
a small open economy assumption. Díaz-Gimenez and Díaz-Saavedra (2017) showed that under the pension system rules governing in 2013, the pension reserve fund would have accumulated a debt of 9% of GDP in 2050. Moreover, the consumption tax necessary to balance the budget would have been 22.6%. Sánchez Martín and Sánchez-Marcos (2010) computed a tax increase need of 5.7 percentage points between 2020 and 2050 to finance the Spanish pension system.

[21] This finding aligns with McGrattan and Rogerson (2008), who observed a similar decrease in male average weekly working hours in the U.S. following women's incorporation into the labor market.

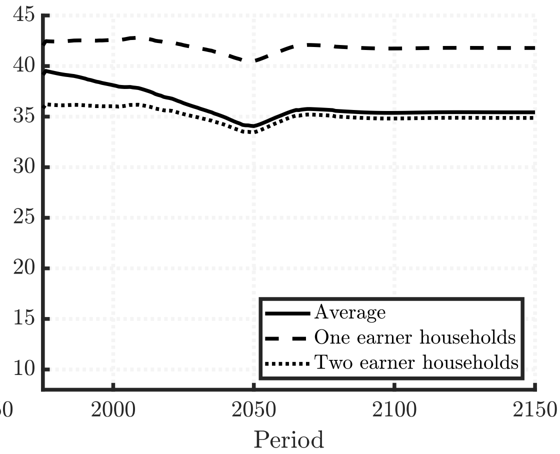
Figure 3.15 Labor Market Outcomes in the Benchmark economy



(a) Share of Modern Households



(b) Female Weekly Hours Worked



(c) Male Weekly Hours Worked

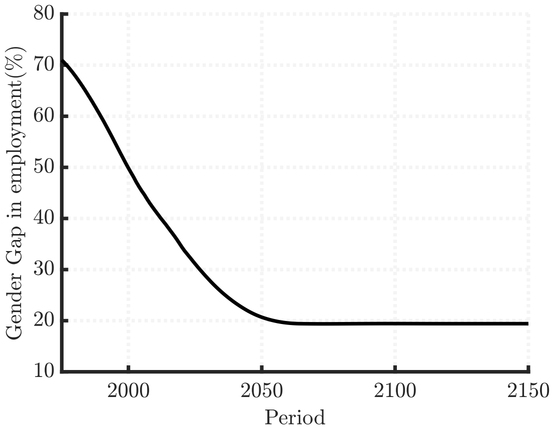
mum levels of 19.4%, 26.3%, and 28.3%, respectively, in the new steady state. Panel (d) presents the gap in average retirement pensions. This gap decreases between the two steady states because women work and contribute more to Social Security and thus have a higher replacement rate. In the new steady state, the gap in the average retirement pension is 15.8%.

Despite the general reductions in the gender gaps, the new steady state in 2100 still exhibits severe gender disparities. In Section 3.8, I explore potential tax reforms aimed at mitigating these gender gaps.

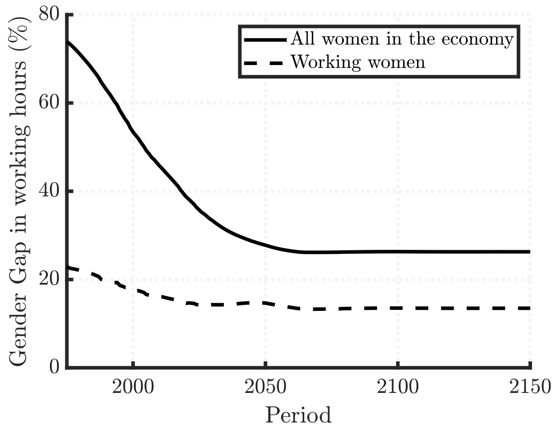
3.7 The Role of Women during the Transition

To quantify the contribution of women to financing pensions during the demographic transition, I construct a novel measure of the share of men's pensions financed by women and vice versa. In this section, I show the evolution of this indicator and discuss the main drivers behind the observed changes.

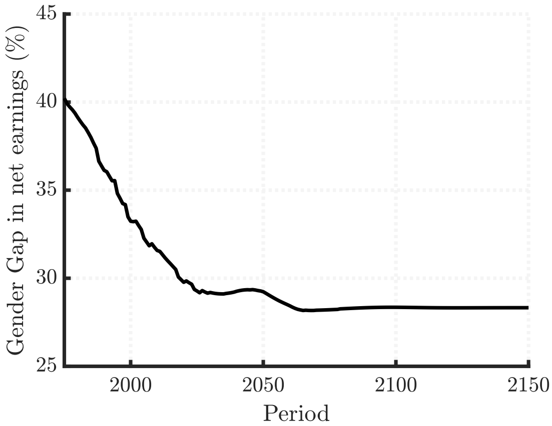
Figure 3.16 Gender gaps in the Benchmark economy



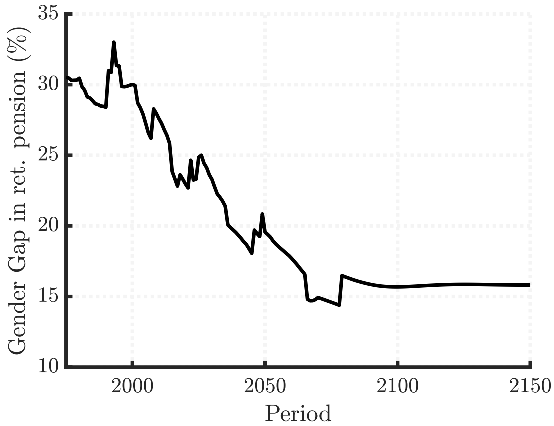
(a) Employment



(b) Average Working Hours



(c) Net Earnings



(d) Retirement Pension

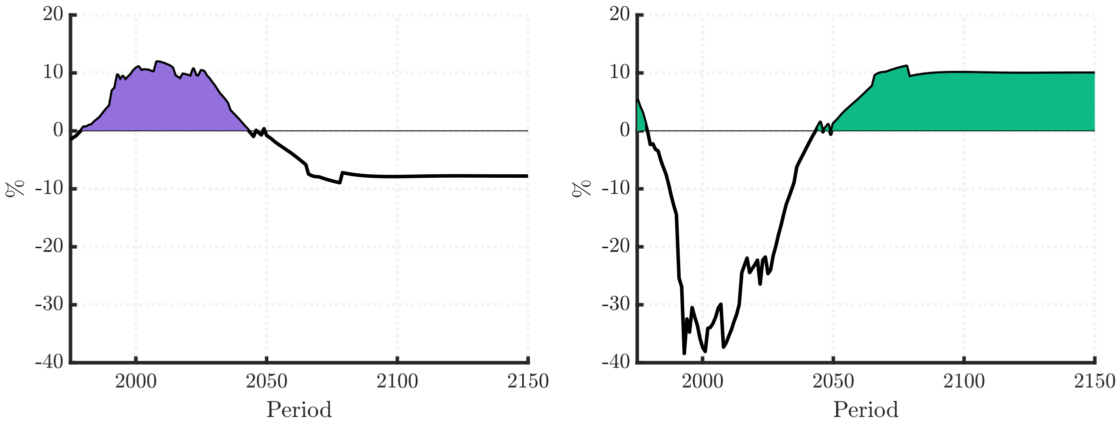
Note: All the gender gaps in this Figure compare working women, women in modern households, to all men in the economy except Figure 3.16b where I compare women in modern households to all women in the economy. Gender gaps are calculated by subtracting the mean of the variable of interest for men from the mean variable of interest for women. Divide the result by the mean variable of interest for men and multiply by 100.

I define the Female Gender Imbalance Indicator (FGII) as the ratio between women's contributions minus women's pensions and men's pensions. Thus, the FGII represents the share of men's pensions financed by women after financing their own retirement and widow pensions. Similarly, the Male Gender Imbalance Indicator (MGII) is the ratio between men's contributions minus men's pensions and women's pensions. The MGII represents the share of women's pensions financed by men after financing their retirement pensions. If the FGII is positive, this implies that women's contributions to Social Security are sufficient to fund their retirement and widow pensions plus a share of men's retirement pensions. In turn, if the MGII is positive, this implies that men's contributions to Social Security cover their retirement pensions plus a share of women's retirement and widow pensions.

Figure 3.17 illustrates the evolution of the FGII and MGII between steady states.

This figure shows that the redistribution of resources between genders is not monotonic during the transition. In 1975, the FGII and the MGII are -1.5% and 5.4%, respectively, implying that women's contributions to Social Security are lower than women's pension expenditures and men finance 5.4% of women's pensions. This is due to widow pensions. In 1975, the contributions women make to Social Security are sufficient to cover the expenditures on women's retirement pensions but not the widow pensions.[22]

FIGURE 3.17 Gender Imbalance Indicator in the Benchmark economy



(a) Female GII (b) Male GII

The FGII increases after 1975. In particular, between 1993 and 2026, women finance, on average, 10% of men's pensions. This is due to several factors. First, women's participation in the labor market increases. Second, on average, women work more per week. Third, the Social Security contributions rate is higher. These three factors generate increasing female contributions to Social Security. Furthermore, in 2019, 36.7% of the pensioners are women, which implies that the extra revenue from women's contributions finances men's pensions.

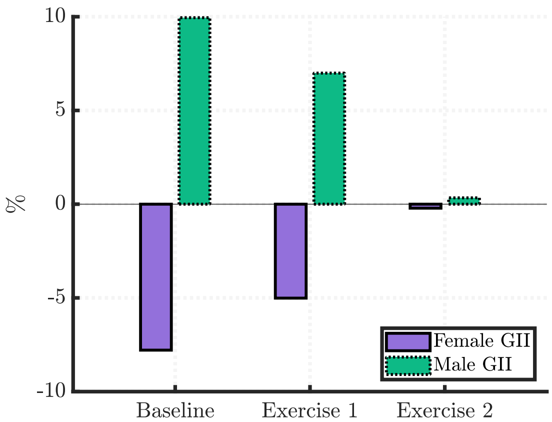
After 2026, the FGII decreases linearly, reaching -7.9% in the new steady state. This implies an MGII of 10%, which means that men are financing 10% of women's pensions. There are several reasons for this. First, more women are entitled to a retirement pension. In particular, in the new steady state, 44.6% of pensioners are women. Second, the average pension for women increases. Third, women live longer than men; thus, ceteris paribus, the pension expenditure for women is greater than for men. Fourth, more women receive both retirement and widow pensions. Finally, women are more likely to be entitled to minimum pensions. In the final steady state, 5.3% and 1.02% of women and men are entitled to a minimum pension, respectively.

I perform a decomposition exercise to understand the role of different factors in

[22]Note that in 1975 only in 38% of the households women participate in the labor market and they work an average of 10.2 hours per week. Figure A3.3 in Appendix illustrates a modification of the FGII, considering only retirement but not widow pensions. This figure shows that in 1975 women finance 2.2% of men's pensions.

generating the fact that women's contributions to Social Security are lower than the expenditure on women's pensions in the final steady state. I simulate the final steady state under two different scenarios. First, I eliminate the existence of bounds in the pension entitlements. Second, I assume that men's and women's life expectancy is the same. This exercise, by construction, eliminates widow pensions. Figure 3.18 compares the FGII and the MGII in the baseline economy, the first and the second exercise, respectively. This Figure shows that if I remove the pension bounds, men finance 2.95 percentage points fewer women's pensions. Thus, pension entitlement limits are not the primary cause of a negative FGII in the new steady state. However, if I set the same life expectancy for men and women, in the new steady state, men finance 0.33% of women's pensions. Therefore, this exercise highlights that women's longer life expectancy and the provision of widow pensions primarily explain the long-term redistribution of resources from men to women.

FIGURE 3.18 Gender Imbalance Indicator Under Two Hypothetical Scenarios



Note: Exercise 1 denotes an economy where I eliminate the bounds in the pension formula. Exercise 2 denotes an economy where I assume equal life expectancy for men and women.

In a nutshell, until 2050, there exists a "women bonus" wherein approximately 10% of men's pensions are financed by women. However, this bonus gradually diminishes over time due to the progressive nature of pension rules, women's longer life expectancy than men, and women's eligibility for widow pensions. As a result, men end up redistributing resources to women. In Appendix A3.5, I provide an additional analysis where I simulate the model economy, assuming that women maintain the same policy functions for employment and hours worked as in the initial steady state. In this exercise, I find that if women did not increase their employment rate and hours worked as they did, the Social Security contributions rate would have risen by 3 to 5 percentage points in 2019. However, the tax rate would have decreased by 0.7 to 0.3 percentage points in the final steady state.

3.8 Gender-Based Taxation in Social Security

The redistribution of resources from women to men in the next decades will be another obstacle preventing women from closing the gender gap in employment and labor supply. Given the redistribution of resources from men to women in the long run, introducing separate Social Security budgets by gender where women and men finance their pensions is a temporary solution. In the long run, women would need to face a higher tax rate than men to finance their pension expenditure, and this will amplify the persistent gender gaps shown in Section 3.6.3.

In this section, I show that Gender-Based Taxation in Social Security considerably narrows the gender gaps in earnings, participation, and hours worked and creates welfare gains for newborn cohorts. Section 3.8.1 discusses its theoretical framework. Section 3.8.2 explains the methodology I use and the main assumptions I make. Lastly, Section 3.8.3 and 3.8.4 show the simulation results and the welfare analysis, respectively.

3.8.1 Theoretical Framework

The tax experiment proposed here relies on a theoretical framework that shows that women should be taxed at a lower rate than men because their work elasticity is lower. Rosen (1977) and Boskin and Sheshinski (1983) pioneered this literature by showing the efficiency gains from taxing women and men at different rates. More recently, Alesina et al. (2011) showed that Gender-Based Taxation (GBT, from now on) with lower taxes for women is superior to the ungendered tax rate. The idea behind this result is that taxing women at a lower rate increases female labor participation and the discrimination cost for employers. At the same time, men's labor supply is more rigid, and they do not reduce their labor supply by much when their marginal tax increases. This tax reform encourages women to increase their labor participation, an explicit goal of the European Union's Lisbon agenda between 2000-2010 and Europe 2020 between 2020-2030 by the European Commission.[23]

3.8.2 Methodology

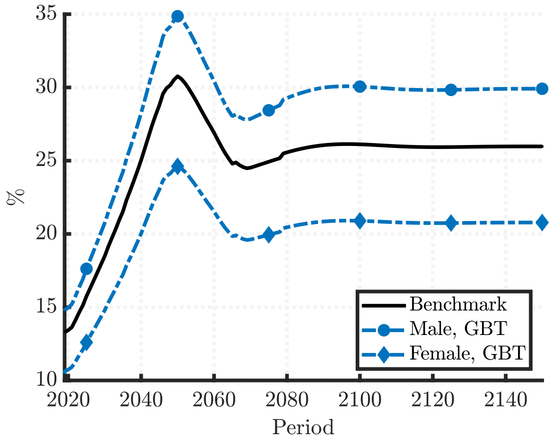
In this policy experiment, the government in 2019 announces that women will always have a lower tax rate than men. In particular, the percentage difference in the tax rate between genders is calculated so that gender imbalances are zero at the introduction of the policy in 2019. As a result, the tax rate on women's earnings is 20% lower than the one at the baseline transition (at every period). Men's Social Security tax rate is adjusted yearly to ensure a balanced budget.

[23]See Bongardt and Torres (2020) for a summary of these two strategies.

3.8.3 Findings

The consequences for the Social Security contributions rate. Figure 3.19 shows the evolution of taxes for men and women. Compared to the Benchmark economy, the tax on men is 1.5 percentage points higher in 2019, while the tax on women is 2.6 percentage points lower. The tax differential between men and women reaches its maximum level in 2050 when the tax on men is 4.1 percentage points higher and the one on women is 6.1 percentage points lower than in the Benchmark economy. These differences in the final steady state are 3.9 and 5.2 for women and men, respectively.[24]

Figure 3.19 Social Security Contributions Rate in the GBT Policy Experiment



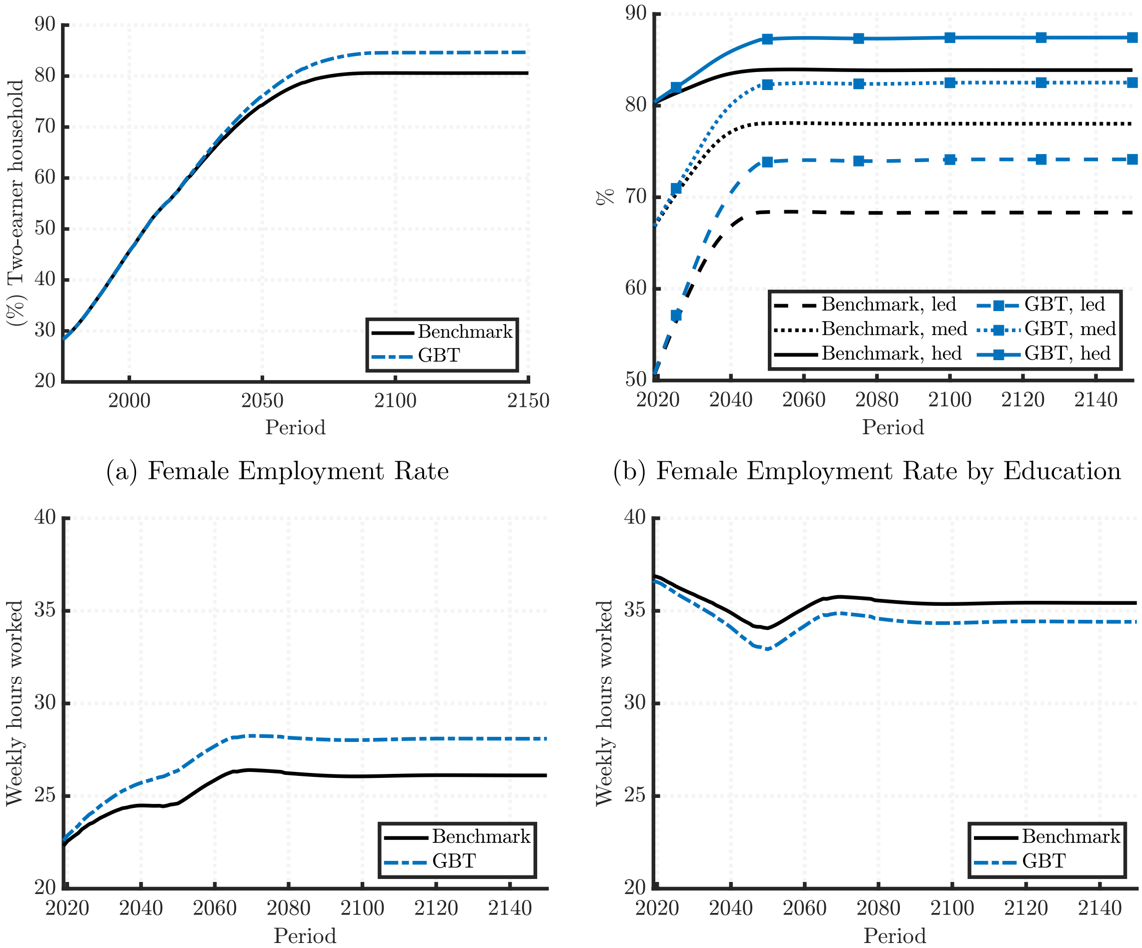
Note: Previous to 2019, the tax rate for women and men is the same and equal to the Benchmark economy.

The consequences for labor supply. Changes in the tax rate levied on women and men significantly impact the labor market. Figure 3.20 compares the evolution of the average female employment rate, female employment rate by education, female labor supply, and male labor supply between the Benchmark and the GBT economy. Panel (a) shows the extensive margin of labor supply for women, i.e., the evolution of the female employment rate. This figure shows that taxing women at a permanently lower rate increases the female employment rate by four percentage points in the final steady state. Panel (b) displays the female employment rate by educational attainment. This figure shows that women with lower education react more to tax changes because their elasticity of labor supply is higher. Panel (c) shows the evolution of the intensive margin of labor supply for women, i.e., the average working hour per week. This figure shows that the average woman works one hour more per week than the average woman in the Benchmark economy. Finally, Panel (d) shows the evolution of the weekly hours worked by men. This figure shows that the average

[24] See Figure A3.4 for an overview of the percentage point tax differences between the Benchmark economy and the GBT policy experiment.

labor supply falls by one hour per week in this economy compared to the Benchmark economy. Thus, this policy disincentives men to work. A greater share of modern couples—in which men supply fewer weekly work hours—and a higher tax rate levied on men explain the drop in men's average working hours.[25]

FIGURE 3.20 Labor Market Reactions by Gender in the GBT Policy Experiment



(a) Female Employment Rate (b) Female Employment Rate by Education

(c) Female Labor Supply (d) Male Labor Supply

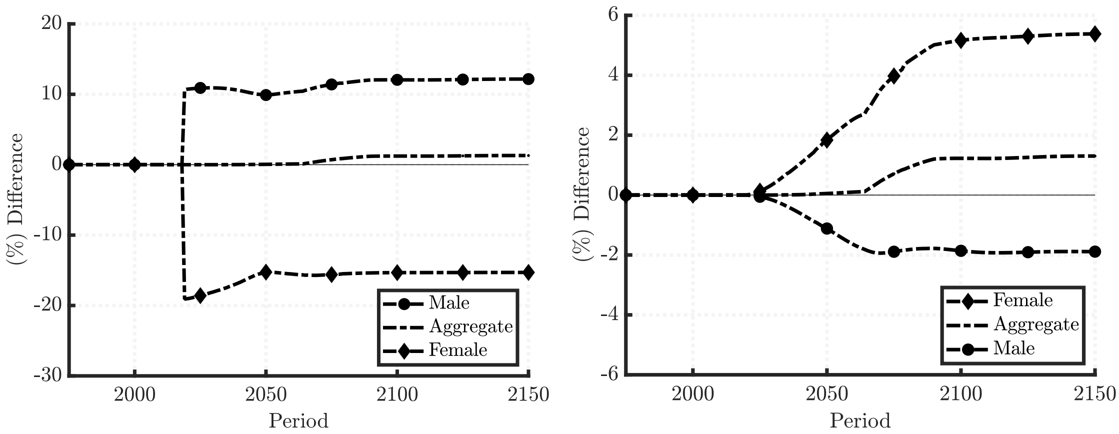
Note: In Panel (b), led, med, and hed represent women with low, medium, and high education, respectively.

The consequences for Social Security expenditures and revenues. Figure 3.21 compares the percentage difference in contributions and pensions by gender and the average difference between the GBT and the Benchmark economies. Panel (a) shows the percentage difference in contributions. It shows that compared to the Benchmark economy, female contributions fall by 15.3% and male contributions increase by 12.16% in the final steady state. This is mostly due to the lower tax rate

[25] According to Figure A3.5 in the Appendix, the shift is attributed to a decrease in average hours worked by modern couples, i.e., there is a substitution effect between genders in those couples where both members have a paid job.

for women and the higher tax rate for men. This tax change is not compensated by the higher female labor supply nor the drop in male hours worked. Panel (b) shows the percentage difference in pension expenditure. It shows that in the final steady state, female pension expenditure rises by 5.4% while men's decreases by 1.9%. In this case, the labor supply reactions of women and men fully explain this result.

Figure 3.21 Percentage Difference of Contributions and Pensions in the GBT Economy Relative to the Benchmark Economy

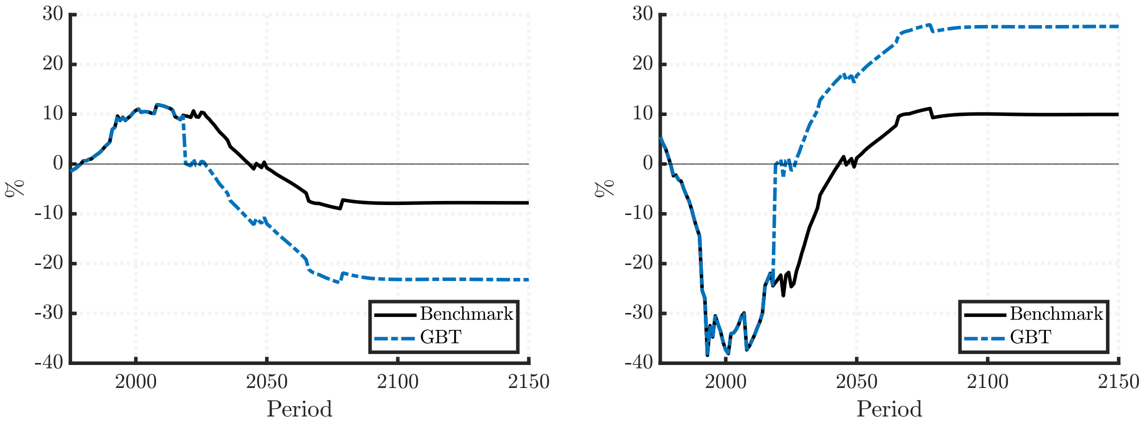


(a) Contributions (b) Pensions

The consequences for the Gender Imbalance Indicator. Figure 3.22 compares the FGII in the left panel and the MGII in the right panel in the Benchmark and GBT economies. By construction, in 2019, the FGII and the MGII are zero because I eliminate the redistribution of resources between genders this year. After 2019, the FGII is negative, and the MGII is positive. In the final steady state, men finance 9.9% and 23.55% of female pensions in the Benchmark and the GBT economy, respectively. The higher female employment rate and hours worked are offset by a lower tax path and a wage gap for women, which explains this result.

The consequences for the gender gaps in the labor market. Although this experiment does not achieve full gender parity, it provides a relevant source to reduce gender pay and working gaps. Figure 3.23 shows the gender gap comparison in hours worked, net earnings, employment, and retirement pension between the Benchmark and the GBT economies. Panel (a) shows the gender gap in average weekly hours worked. This figure indicates that the gap reduces by 10.7%. This drop is not solely due to women providing more work per week but also to the fact that men in modern couples are reducing their labor supply. Panel (b) shows the gender gap in average net earnings. This figure indicates that the gap in net earnings between men and women decreases by 5.9%. This decrease is driven by the tax difference by gender in this economy compared to the Benchmark economy. Panel (c) shows the gap in the extensive margin of labor supply. This figure shows that gender

Figure 3.22 Gender Imbalance Indicator in the GBT Policy Experiment



(a) Female GII (b) Male GII

Note: The Female GII is the ratio of female contributions minus female pensions to male pensions. The Male GII is calculated as the male surplus over female pensions.

gaps in employment decrease by 4.07 percentage points. Finally, Panel (d) shows the gender gap in the average retirement pension. In the GBT economy, a higher share of women is eligible for a retirement pension due to their long working hours and greater net earnings, which translates to a 4.1 percentage points reduction in average pensions compared to the Benchmark economy.

3.8.4 Welfare Analysis

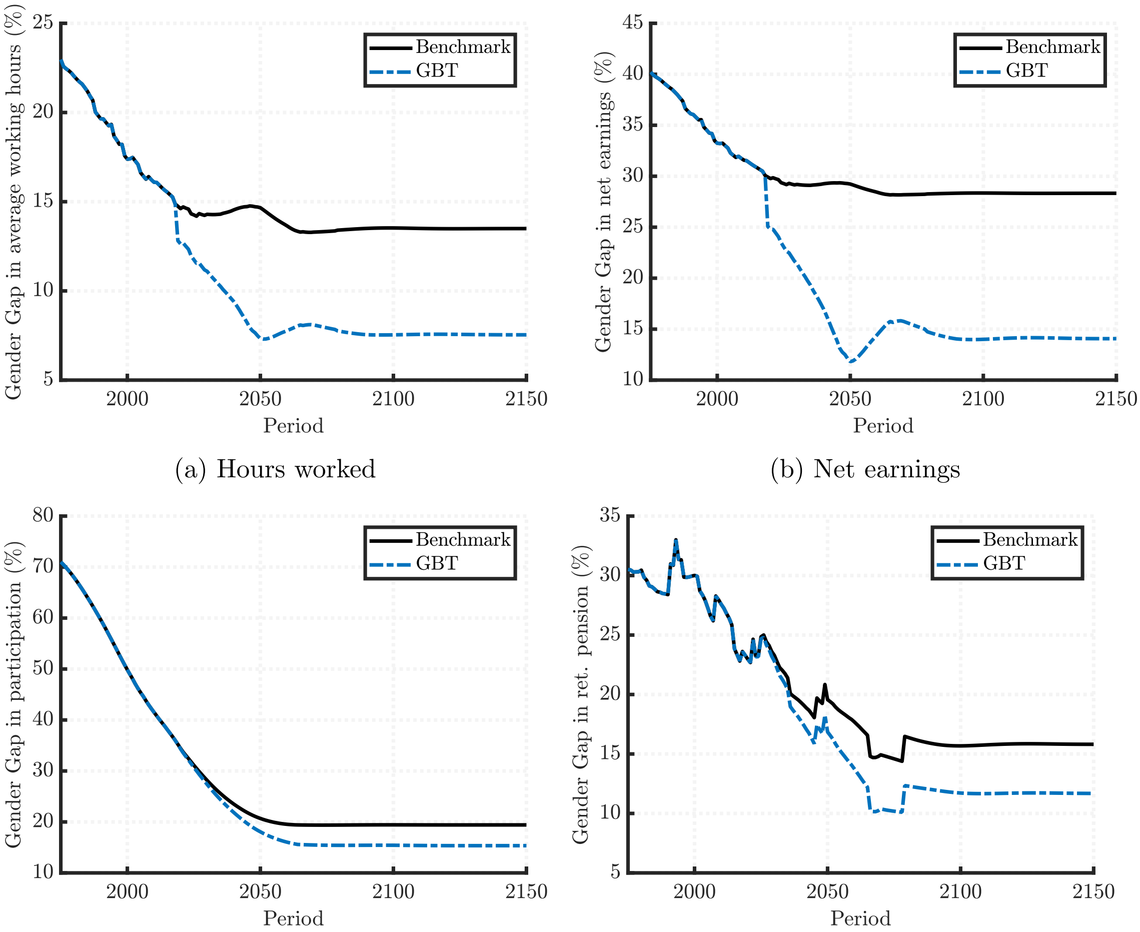
Methodology. I measure welfare by computing the per-period consumption compensation under the Benchmark economy, which guarantees equal discounted welfare under the current tax system in 2019 and the one implied by the tax reform (GBT). Then, the welfare gains or losses of cohort s are calculated as the constant percentage change, λ_s, in the baseline consumption path, such that these households are indifferent between the current tax system and the tax reform. Thus λ_s is given by

$$\lambda_s = \left(e^{\mathbb{E}V\left(\ddot{\boldsymbol{\Psi}}_{\mathbf{s}}, \ddot{\boldsymbol{\Upsilon}}_{\mathbf{t}}; \ddot{\boldsymbol{\Omega}}_{\mathbf{t}}\right) - \mathbb{E}V\left(\boldsymbol{\Psi}_{\mathbf{s}}, \boldsymbol{\Upsilon}_{\mathbf{t}}; \boldsymbol{\Omega}_{\mathbf{t}}\right)} - 1 \right) \times 100, \tag{3.10}$$

where $\boldsymbol{\Psi}_{\mathbf{s}}$ denotes the state variables of the household, $\boldsymbol{\Upsilon}_{\mathbf{t}}$ the aggregate state variables of the economy at time t and $\boldsymbol{\Omega}_{\mathbf{t}}$ is the government policy schedule regarding the Social Security tax rate at time t. All the variables with the accent, $\ddot{x}$, refer to the states and tax policy under the GBT tax reform.

The interpretation of this welfare measure is as follows: λ_s denotes the percentage of consumption cohort s needs to receive to be indifferent between the Benchmark economy and the GBT economy. When $\lambda_s > 0$, it implies that household aged j at time t would be better under the reform scenario. In contrast, this household would

Figure 3.23 Gender Gaps in the Benchmark Economy and the GBT Policy Experiment



(a) Hours worked

(b) Net earnings

(c) Employment

(d) Retirement pension

like to remain under the Benchmark economy scenario if $\lambda_s < 0$.

Main assumptions. First, these households decide to form a traditional or a modern household at the beginning of their life cycle. When making this decision, they draw the participation cost for women, and they have perfect foresight about the path of taxes, i.e., the one of the Benchmark economy. If the discounted expected value of forming a modern household (net of participation costs) outweighs the discounted expected value of forming a traditional couple, this household chooses to form a modern couple. Otherwise, it forms a traditional couple. Second, the shift to GBT is taken as an unanticipated shock. All cohorts remain traditional or modern in 2019 based on their decision when they are born. Essentially, I keep couples from reoptimizing their household type decision. It is possible, however, for them to adjust their labor supply, consumption, and savings in this new economy.[26]

[26]This assumption is consistent with the baseline model, in which individuals only choose their household type at the beginning of their lives and cannot reoptimize it.

Cohorts alive in 2019. Table 3.6 summarizes the average consumption compensation for different household types alive in 2019. While, on average, households prefer to remain in the Benchmark economy (-0.1% of lifetime consumption), there are substantial heterogeneous effects among households. Specifically, traditional couples suffer welfare losses of 1.49% of their lifetime consumption compared with modern couples, with welfare gains of 0.25%.

Traditional couples suffer welfare losses because they cannot change their household composition, and men's taxes are higher than in the Benchmark economy. Interpreting the welfare gains for modern couples is less straightforward. This is because the Gender-Based Taxation economy increases the husband's tax rate and decreases the wife's. This analysis shows that the latter effect outweighs the former, and therefore modern couples are better off under the new tax system.

Table 3.6 Welfare Effects of the GBT Policy Experiment for Cohorts Alive in 2019

Type	**% Consumption Compensation**
All households	-0.10
Traditional households	-1.49
Modern households	0.25
Low educated female	-0.44
Medium educated female	-0.19
Highly educated female	0.12

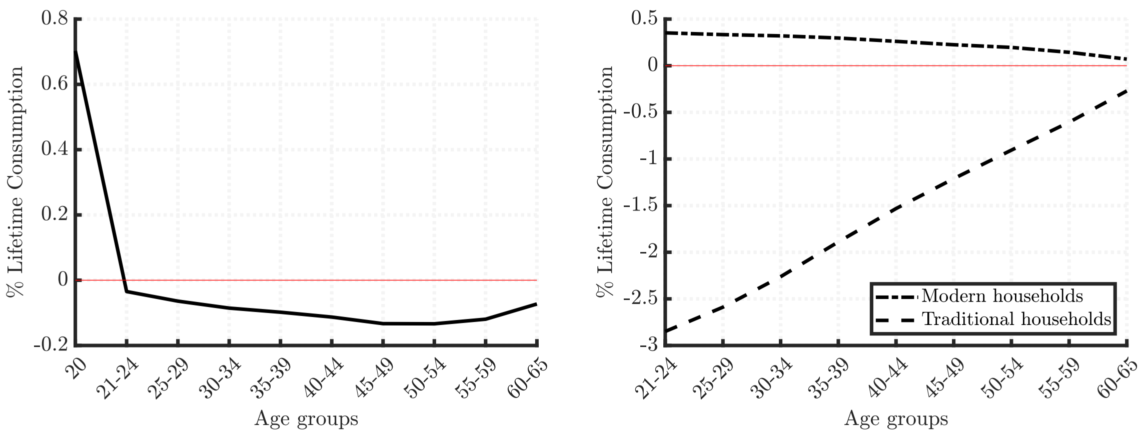
Table 3.6 also shows that welfare increases with the wife's education. Women with low education experience a welfare loss of -0.44% lifetime consumption, whereas women with college degrees experience a welfare gain of 0.12%. The effect is purely compositional: women with low education work less and are more likely to form traditional marriages. Table 3.7 adds to the previous analysis by assessing the welfare implications by spousal education. It shows that welfare losses are significantly higher among households with low- and medium-educated men but decrease with women's education. One interesting result is that households with a highly educated wife and a low or medium-educated husband receive welfare gains since women are the primary income providers in their households. Finally, Figure 3.24 shows a

Table 3.7 Welfare Effects of the GBT Policy Experiment by Spouses' Education for Cohorts Alive in 2019

		Female		
		Low educated	Medium educated	Highly educated
Male	Low educated	-0.36	-0.01	0.39
	Medium educated	-0.52	-0.18	0.23
	Highly educated	-0.71	-0.38	0.03

decomposition exercise of lifetime consumption compensation by age group (3.24a) and age group and household type (3.24b). The first result in 3.24a is that although only newborn cohorts are better off under this tax reform scenario, the losses of cohorts alive in 2019 vary with age. This chart displays a U shape, indicating that the eldest and youngest cohorts have lower welfare losses than the middle cohorts. Between 21 and 54, welfare drops because the older the cohort, the greater the share of traditional households. Welfare increases in cohorts older than 54 because the policy change affects only a few working years, so losses are minimal. The second result is in line with this last finding. Figure 3.24b shows that welfare losses and gains by household type decrease with age. This is explained by older cohorts living under the new policy regime for fewer years than younger cohorts for whom the changes persist for longer.

Figure 3.24 Welfare Implications by Age Groups and Household Type of the GBT Policy Experiment



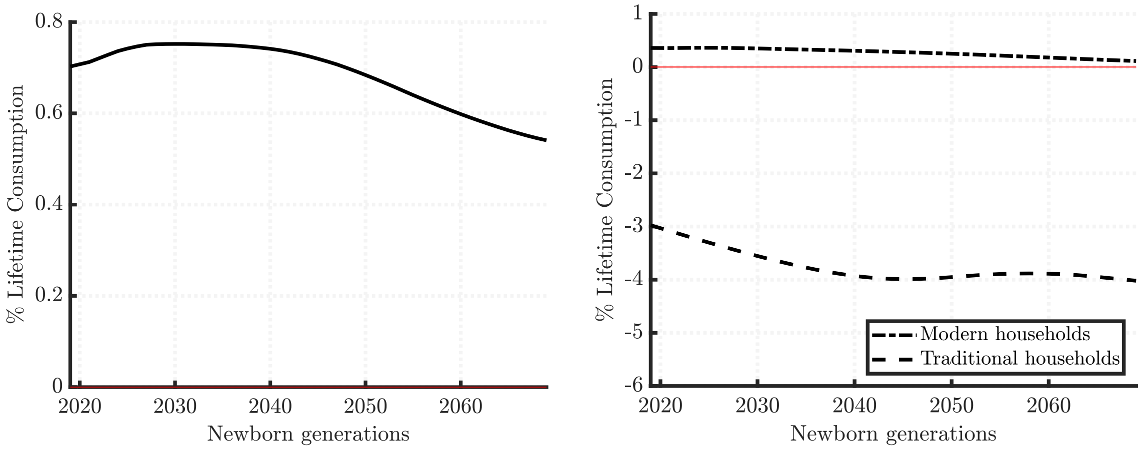
(a) All households

(b) By household type

Note: Traditional households refer to one-earner households where only the husband works, while in modern households, both partners are in the labor market.

Newborn cohorts. Figure 3.25 presents the percentage of lifetime consumption for newborn cohorts on average (3.25a) and by household type (3.25b). Panel (a) shows that newborn cohorts are, on average, better off. Despite this, welfare gains range from a maximum of 0.75% in 2027 to 0.55% for the younger cohorts. Panel (b) shows that modern households have welfare gains and traditional households have welfare losses, but the magnitudes vary by cohort. In particular, modern couples' welfare gains decrease for younger cohorts while traditional couples' welfare losses increase for younger cohorts. Tax differentials between men and women in this policy experiment and the Benchmark economy explain these differences by cohort. Figure A3.4 in the Appendix shows that these differences increase yearly, achieving a maximum in 2050 and dropping after that.

Figure 3.25 Welfare Implications for Newborn Cohorts of the GBT Policy Experiment



(a) All households (b) By household type

Note: Traditional households refer to one-earner households where only the husband works, while in modern households, both partners are in the labor market.

The introduction of Gender-Based Taxation (GBT) leads to an overall increase in utility for modern households. The augmented utility stems from higher levels of consumption and increased leisure time. Although there are slight variations in gross earnings between GBT and the Baseline economy, net earnings are higher under GBT. This can be attributed to the greater elasticity of labor supply among women compared to men. Consequently, women tend to respond more strongly to tax changes, leading to a higher employment rate than men's reduction. Regarding leisure time, both the husband's and wife's leisure contribute to the household's utility function with the same weight, but childcare responsibilities reduce the wife's leisure. Due to the decrease in labor supply, men derive more benefits from leisure, which ultimately outweighs the reduction in leisure by women.

Finally, regarding the education of newborn cohorts, two observations are relevant. First, in modern households, welfare gains and losses are heterogeneous across households with different levels of education between the wife and husband. On the one hand, those whose wives are more educated than their husbands benefit more from the tax reform. For instance, in 2019, modern households consisting of a highly educated woman married to a low-educated man would have to receive 0.9% in consumption to compensate for being in the Benchmark economy instead of the GBT Experiment. On the other hand, those households where the husband's education is greater than the wife's face the highest welfare losses as men are the primary income providers. For example, in 2019, modern households composed of highly-educated husbands and low-educated wives should receive -0.32% in consumption compensation for the GBT Experiment. Second, for traditional households, welfare is homogeneous by education because the hours worked do not change dramatically by the husband's education.[27]

Overall, the Gender-Based Taxation Experiment narrows gender gaps in the labor market, and new generations benefit from it in terms of welfare.

3.9 Conclusion

While the increase in female labor force participation has been largely studied, this is the first paper to quantify the effect of this increase on Social Security finances. Using a deterministic overlapping generations model of married households, I find that women are crucial to finance the Spanish pension system until 2050. Women's contributions to Social Security are higher than their pension expenditure, and therefore, they finance a positive share of men's pensions. In particular, between 1993 and 2026, women finance 10% of men's pensions. Moreover, the rise in female employment rate and hours worked prevents Social Security from increasing the tax rate between 1 and 2 percentage points. After 2050, men redistribute resources to women because women live longer than men and receive widow pensions, which are complementary to retirement pensions by law.

I use the model as a laboratory to quantify the effects of introducing Gender Based Taxation in the Social Security system. In my experiment, in 2019, the government announces that the tax rate for women is 20% lower than the tax path of the baseline economy, and the men's tax rate adjusts to balance the budget. Although this experiment does not achieve gender parity, introducing Gender-Based Taxation reduces gender differences in average employment, earnings, and retirement pensions. At the same time, it displays generalized welfare gains of 0.55-0.7% in consumption compensation. Therefore, gender-sensitive strategies are critical to prevent long-term scarring in the labor market.

My paper provides numerous contributions. First, it delivers a quantitative answer to a question not previously addressed with a model replicating key features of the Spanish economy between 1975 and 2019. Second, it carefully models the Spanish Social Security System and captures the central transitional dynamics in Spain during the last forty-five years. Finally, it offers an excellent tool for addressing different policies, such as eliminating widow pensions or taxing consumption to finance Social Security extra costs without increasing its rate or analyzing previous Social Security reforms implemented in Spain.

There are several caveats to be noted regarding the present study. First, labor force participation for women is a one-time decision at the beginning of their lives. As a result, women out of the labor force cannot react to policy changes by

[27] Tables A3.5 and A3.6 in the Appendix display the winners and the losses of newborn generations by education for modern and traditional households, respectively.

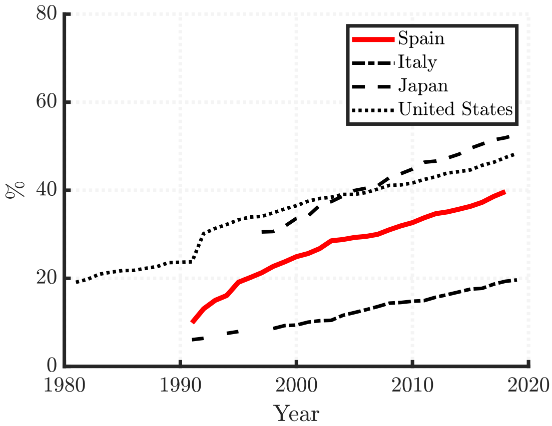
reconsidering their labor supply decisions. Second, this paper only models married individuals, leaving a significant share of the population out of the analysis. However, I believe most single individuals self-finance their retirement pensions, and therefore the conclusions of this paper will not be altered.

Further research may include a general equilibrium analysis of singles and married couples with endogenous fertility and retirement decisions. Additionally, a data comparison of the gender imbalances measure I propose in this paper is left for future research. Finally, further cross-country comparison research is needed to fully understand the implications of the increase in women's employment rate on pension funding.

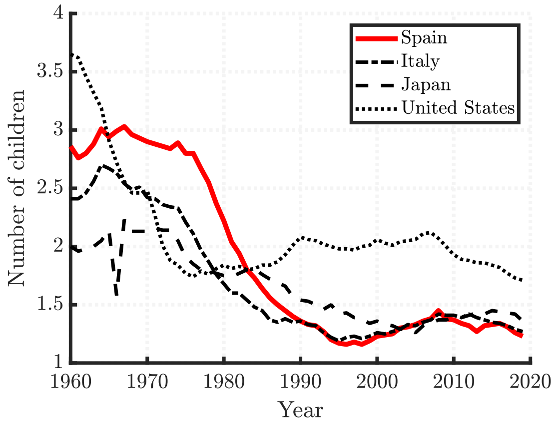
A3 Appendix to Chapter 3

A3.1 Figures and Tables

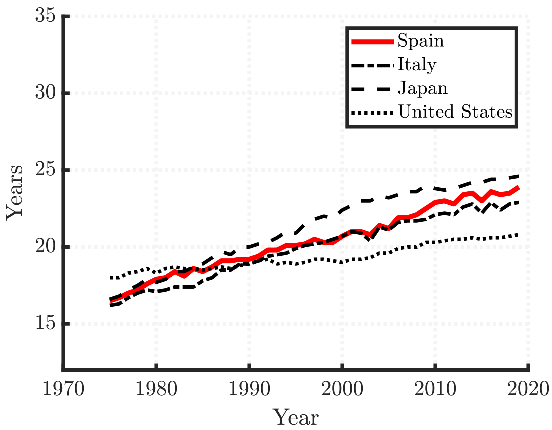
Figure A3.1 Demographic Trends in Selected Countries



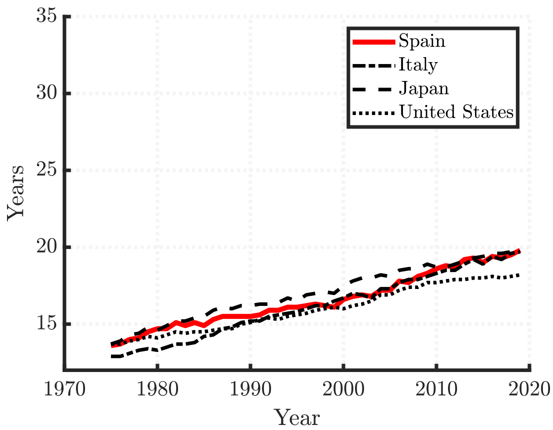
(a) Share of Population With a Tertiary Education



(b) Total Fertility Rate



(c) Female Life Expectancy at 65



(d) Male Life Expectancy at 65

Source: OECD statistics

Figure A3.2 Model Fit: Average Lifetime by Year

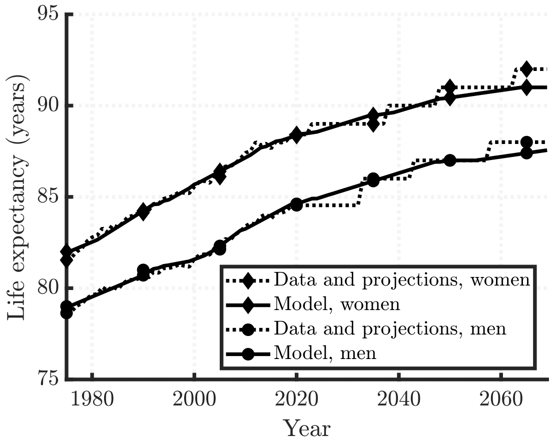
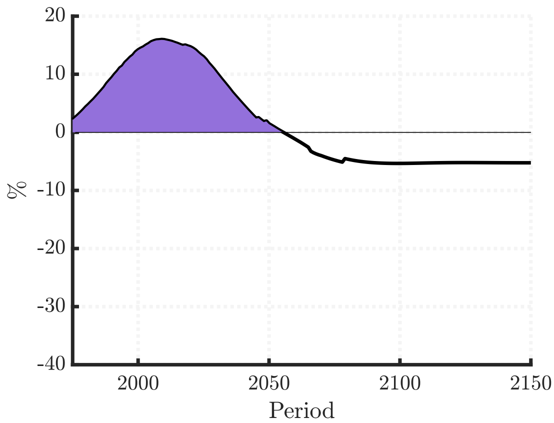


Figure A3.3 Modified Female Gender Imbalance Indicator in the Benchmark economy



Note: The Modified Female Gender Imbalance Indicator is the ratio of female contributions minus female retirement pensions to male pensions.

Figure A3.4 Percentage Difference in the Social Security Contributions Rate by Gender Between the Benchmark and the GBT Economies

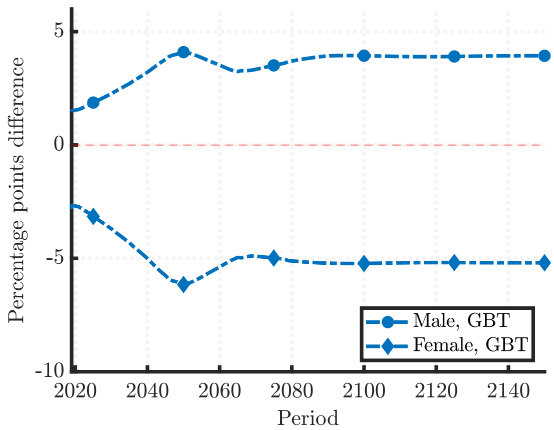
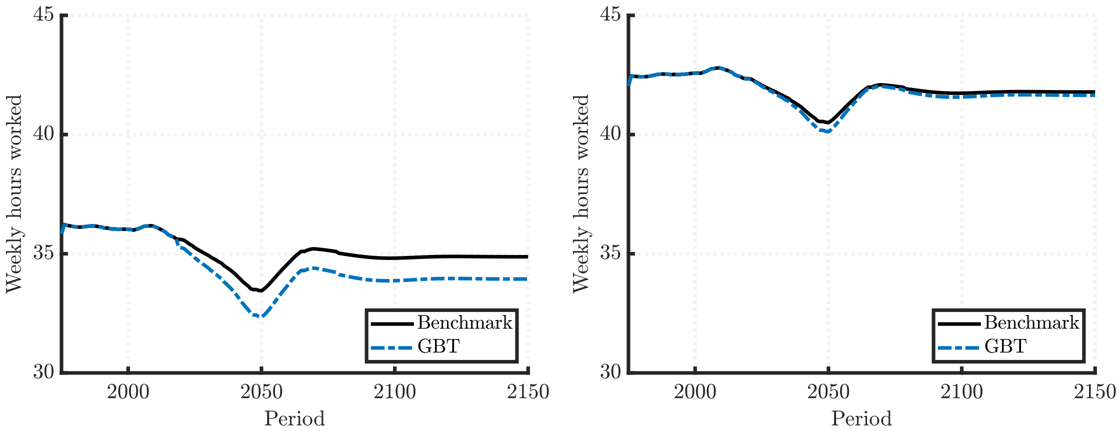


Figure A3.5 Men's Average Worked Household by Household Type in the Benchmark and the GBT Economies



(a) Modern household (b) Traditional household

Table A3.1 Key Parameter Changes of Old-Age Pension Reforms in Spain Since 1980

Reform	Legal retirement age	N_c	N_f	N_b
Prior 1985	65	8	35	2
1985	65	15	35	8
1997	65	15	35	15
2002	65	15	35	15
2007	65	15	35	15
2011	67*	15	37	15**
2013	67*	15	37	16**

Note: * or 65 if 38.5 years of contributions. ** 25 years from 2022 onwards. N_c, N_f, and N_b represent the years of contributions for pension eligibility, for a full pension, and for the pension calculation, respectively.

Table A3.2 Calibration: Initial Steady State

	Parameter	Value
Demographics		
Husband's life expectancy (periods)	J_h	60
Wife's life expectancy (periods)	J_w	63
Women's mean age of first child	J_c	6
Fixed time cost of babies	h	0.046
Fertility rates	$\phi(1, \bar{s})$	2.5
	$\phi(2, \bar{s})$	2.2
	$\phi(3, \bar{s})$	1.8
Household preferences		
Intertemporal elasticity of leisure	γ	4
Discount factor of households	β	0.992
Taste of leisure	θ	0.58
Distribution participation cost		
Shape low educated	$\varphi(1, \bar{s})$	9.9
Shape medium educated	$\varphi(2, \bar{s})$	7.9
Shape high educated	$\varphi(3, \bar{s})$	5
Pension		
Retirement age	J_R	46
Elegibility (years of experience)	N_c	15
Contributed years	N_b	15
Survivor pension: share of husband's pension	χ	52%
Minimum retirement pension with dependent	$\underline{P}$	29%
Minimum retirement pension without dependent		24%
Minimum disability pension	$\underline{P}d$	23
Maximum pension (both types)	$\bar{P}_t$ and $\bar{Pd}_t$	125%
Penalties pension formula		
Penalties	α_0^p	0.5
	$\triangle\alpha_1^p$	0.03
	α_2^p	0.8
	$\triangle\alpha_3^p$	0.02
	a_1	15
	a_2	25
	a_3	35
Annual interest rate	r	0.036
Annual productivity growth	g	0.0177

TABLE A3.3 Distribution of Pairings (%) by Partners' Level of Educational Attainment and Female Birth Cohort

			Wife's education		
	1910-1919		**Low**	**Medium**	**High**
		Low	80.60%	1.33%	0.34%
		Medium	5.58%	7.14%	0.48%
		High	1.79%	1.72%	1.02%
	1920-1929				
		Low	77.98%	1.63%	0.27%
		Medium	6.14%	9.04%	0.48%
		High	1.42%	2.01%	1.02%
	1930-1939				
		Low	70.25%	2.35%	0.30%
Husband's education		**Medium**	7.77%	12.70%	0.73%
		High	1.45%	2.78%	1.67%
	1940-1949				
		Low	47.73%	4.56%	0.37%
		Medium	10.89%	24.08%	1.60%
		High	1.43%	5.15%	4.20%
	1950-1959				
		Low	17.49%	8.19%	0.52%
		Medium	9.69%	41.72%	4.87%
		High	0.85%	7.28%	9.40%
	1960-1969				
		Low	8.06%	7.52%	0.49%
		Medium	6.04%	49.00%	9.21%
		High	0.32%	6.88%	12.48%
	1970-1979				
		Low	4.29%	5.56%	0.71%
		Medium	3.59%	46.80%	14.80%
		High	0.30%	6.40%	17.55%
	1980-1984				
		Low	1.7%	3.0%	0.5%
		Medium	2.5%	44.8%	16.8%
		High	0.2%	6.7%	23.8%
	1985-1989				
		Low	1.5%	3.5%	0.6%
		Medium	2.5%	41.9%	16.8%
		High	0.2%	6.6%	26.5%
	1990-1995				
		Low	1.4%	2.8%	0.5%
		Medium	1.9%	37.2%	16.8%
		High	0.2%	6.9%	32.2%

Table A3.4 Evolution of Household's Composition

Male & Female	1975	2019	2050	2150
1 & 1	59%	12%	2%	1%
1 & 2	4%	5%	4%	3%
1 & 3	1%	1%	1%	1%
2 & 1	7%	5%	3%	2%
2 & 2	19%	40%	40%	37%
2 & 3	2%	13%	19%	20%
3 & 1	1%	0%	0%	0%
3 & 2	4%	6%	7%	7%
3 & 3	4%	17%	25%	29%

Note: 1 refers to low education, 2 to medium education, and 3 to high education.

Table A3.5 % Consumption Compensation for Newborn Cohorts in Modern Households for the GBT Policy Reform

		Female		
		Low educated	Medium educated	Highly educated
Year 2019				
Male	Low educated	0.20	0.54	0.90
	Medium educated	-0.07	0.28	0.63
	Highly educated	-0.32	0.01	0.36
Year 2039				
Male	Low educated	0.12	0.47	0.85
	Medium educated	-0.13	0.21	0.58
	Highly educated	-0.41	-0.08	0.30
Year 2069				
Male	Low educated	-0.04	0.32	0.71
	Medium educated	-0.30	0.05	0.43
	Highly educated	-0.59	-0.24	0.14

Table A3.6 % Consumption Compensation for Newborn Cohorts in Traditional Households for the GBT Policy Reform

	Male		
Year	Low educated	Medium educated	Highly educated
2019	-2.54	-2.56	-2.58
2039	-2.78	-2.80	-2.84
2069	-3.03	-3.05	-3.09

A3.2 Brief Literature Review on Spanish Pension Reforms

The impact of pension reforms in Spain has been extensively studied in the economic literature through the lens of overlapping generations models and accounting frameworks. Bonin et al. (2001) used a Generational Accounting framework and found that the new legal setting introduced in 1997 could result in future generations facing liabilities as high as 176% of 1996 GDP. Martín (2010) used a heterogeneous-agent dynamic general equilibrium model to examine the pension reforms in 1997 and 2001 and found that they increased the pension system's liabilities. Sánchez Martín and Sánchez-Marcos (2010) used a general equilibrium model where households consist of two potential earners to study the 1997 and 2002 pension reforms and obtained similar results. De La Fuente and Doménech (2013) estimated the effects of the 2011 pension reform with Aggregate Accounting and found that these measures would reduce expenditure by approximately 1.4 percentage points of the GDP. Finally, Díaz-Gimenez and Díaz-Saavedra (2017) used an overlapping generations model with endogenous retirement to analyze the 2001-2013 pension reform and found that while it improved the pension system's sustainability, it came at the expense of a reduction in the real value of the average pension.

A3.3 Aggregating Pensions and Contributions

I assume the Social Security budget is balanced at all periods t. Therefore, the Social Security contributions rates, τ_t, satisfy the following condition:

$$\tau_t = \frac{P_t}{T_t},$$

where P_t represents the aggregate expenditure on retirement and widow pensions at time t, and T_t are the Social Security contributions collected at time t. I describe below how these two aggregate variables are composed.

$$P_t = \sum_{i=\{w,h\}} \sum_{j=J_R}^{J} \sum_{z=1}^{Z} \sum_{x=1}^{X} \sum_{\epsilon^h(z)}^{E(z)} \sum_{\epsilon^w(x)}^{E(x)} \sum_{m=0}^{1} (p^i_{t,j,z,x,\epsilon^h,\epsilon^w,m} + pd_{t,j,z,x,\epsilon^h,\epsilon^w,m})$$
$$M_t(z,x) S\left(\epsilon^h, \epsilon^w\right) F_{t,j}(z,x,\epsilon^h,\epsilon^w) N_{t,j},$$

$$T_t = \sum_{i=\{w,h\}} \sum_{j=1}^{J_R-1} \sum_{z=1}^{Z} \sum_{x=1}^{X} \sum_{\epsilon^h(z)}^{E(z)} \sum_{\epsilon^w(x)}^{E(x)} \sum_{m=0}^{1} l^i_{t,j,z,x,\epsilon^h,\epsilon^w,m} y^i_{t,j,z,x,\epsilon^h,\epsilon^w,m}$$
$$M_t(z,x) S\left(\epsilon^h, \epsilon^w\right) F_{t,j}(z,x,\epsilon^h,\epsilon^w) N_{t,j},$$

where $N_{t,j}$ is the share of individuals of age j at time t and $F_{t,j}(z,x,\epsilon^h,\epsilon^w)$ denotes the distribution of households depending on whether they are traditional or modern.

A3.4 Calibration Details

Earnings

The central database I use to calibrate the earnings is the Continuous Sample of Working Histories *(Muestra Continua de Vidas Laborales, MCVL)*. With the editions 2006-2017, I construct a yearly panel. I restrict my sample to full-time workers. These workers belong to the general regime. I exclude part-time workers and individuals with non-positive wages to work with a stable earnings measure over time.

Distribution of the initial productivity. The first available year in the MCVL is 1980; however, this year's number of observations for highly educated men and women is low. To overcome this issue, I computed the initial productivity by education using the sample from 2001.

Once I estimate those moments of the log-normal distribution for 2001, I make two assumptions to get the distribution in 1975. First, I assume the variance is the same as the one estimated in 2001. Second, I subtracted from the mean of the log-normal distribution the annual growth of earnings between 1980-2001. I computed the average earnings per hour growth between 1980 and 2001 with the Industry and Services Wage Survey.[28] The calculations show that the annual growth was 7, 4%. Table A3.7 summarizes the male's initial earnings distribution parameters. I discretized the initial steady state distribution by considering seven initial productivity and forty-nine possible combinations for each household's educational type.

Table A3.7 Moments of the Log-Normal Distribution

	μ^h_{01}	μ^h_{75}	$\sigma_{01} = \sigma_{75}$	# obs
Led	9.62	9.42	0.327	23453
Med	9.76	9.56	0.396	12971
Hed	9.96	9.76	0.450	7897

Note: 01 accounts for 2001 while 75 does for 1975.

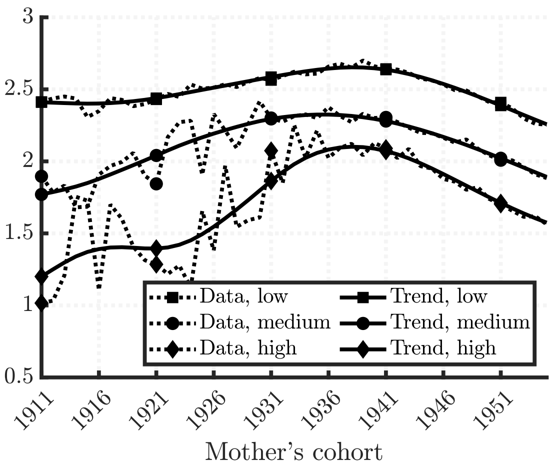
Number of children

To compute the average number of children by educational attainment, I use Zeman et al. (2014) database. The authors of this database extracted micro-data from the Spanish Population Census of 1991 and 2011 provided by the INE. I considered the first census for cohorts born before 1931 and the second for the others. Then, I

[28]INE carried out this survey until 2000, when the quarterly labor cost survey replaced it. I computed this earnings growth per hour as an average of all sectors. I restricted my attention to workers and job categories.

smoothed the series with the Hodrick-Prescott filter. Figure A3.6 shows the cumulative fertility rates by mother's cohort and educational attainment implied by the data and the trend. For the initial steady state, I take the average of the trend. Therefore the three parameters regarding the number of children are $\phi(1) = 2.5$, $\phi(2) = 2.2$, and $\phi(3) = 1.8$ where 1, 2, and 3 denote the mother's education level. These refer to low, medium, and high education, respectively.

FIGURE A3.6 Cumulative Fertility Rate by Education



Childcare cost

Regarding childcare costs, children are only costly when they are babies. I include childcare costs for the first three years of a child's life (0-2). The main reasons for this assumption are two. On the one hand, childcare expenditures for children over three years old might not be a key determinant for women's work. By the school year 2000-01, 89.7% of children aged three were enrolled in schools, and 68.4% of them were in free-of-charge public institutions. On the other hand, the cost is relevant for children younger than three, as few children have access to public kindergartens, which are cheaper but not free-of-charge administrations.

Full-time and part-time workers

Workers in the model are classified into full-time and part-time workers. The Spanish Statistical National Institute (*INE*) defines a full-time worker as an individual that works at least 30 hours per week and a part-time worker if he/she works less than 35 hours per week. Since between 30 hours and 35 hours, a worker can be considered a full-time or part-time worker; I take the midpoint. If the individual works more than 32.5 hours per week, he is considered a full-time worker. Otherwise, it is considered a part-time worker. In the model, a full-time worker is an individual with no less than 0.325 units of labor supply. Additionally, an individual is classified as a participant in the labor market if her labor supply is at least one hour per week, i.e.: $l_t \geq 0.01$.

Distribution of households by education of each spouse

Individuals' education levels are classified as low, medium, and high. Individuals with a low level of education have not completed secondary school; individuals with a medium level of education have not completed a university degree; and individuals with a high level of education have at least completed tertiary education.

In the initial steady state. To compute the initial distribution of households by the educational attainment of each spouse $M_{t=1}(z,x)$ I follow Esteve and Cortina (2006). In particular, I consider the educational attainment of couples and spouses who are residents in principal houses in the Spanish Census of 2001 and 2011. With this information, I computed the share of households for each combination of spouse's education for the 1910-1955 women's cohorts. Table A3.3 represents the data I extracted for cohort groups of 10 years. Nevertheless, I linearly approximate it to get it by cohort. Then, the newborn generation in the initial steady state, $j = 1$, corresponds with the data counterpart for the 1955 cohort. The generation of age $j = 2$ is assigned to the data counterpart in 1954, etcetera.

During the transition. The distribution of households for newborn cohorts follows the same methodology and data as the initial steady state until the cohort of 1979 (corresponding to newborns in the model year 1999). There is no available data for younger cohorts on the educational attainment of couples who are residents in principal households in the Spanish Census of 2001 and 2011. To overcome this challenge, I proceed as follows: 1) I construct a likelihood modifier using the Spanish Census data on education levels of married individuals of generation 1970-79. This is the ratio of the proportion of households of a given type to the proportion implied by random matching between education types. 2) Using the INE data of education of men and women for cohorts 1979-1995 (newborns in years 1999-2015 in the model), I compute a matrix of education levels of couples that would be implied by random matching (that is, just the product of the fraction of individuals of a given gender that have an education level). Then, I multiply this matrix of couples' education levels randomly matched by the likelihood modifier factor to estimate married individuals' educational attainment. See Table A3.3 for the distribution of couples by education and cohort.

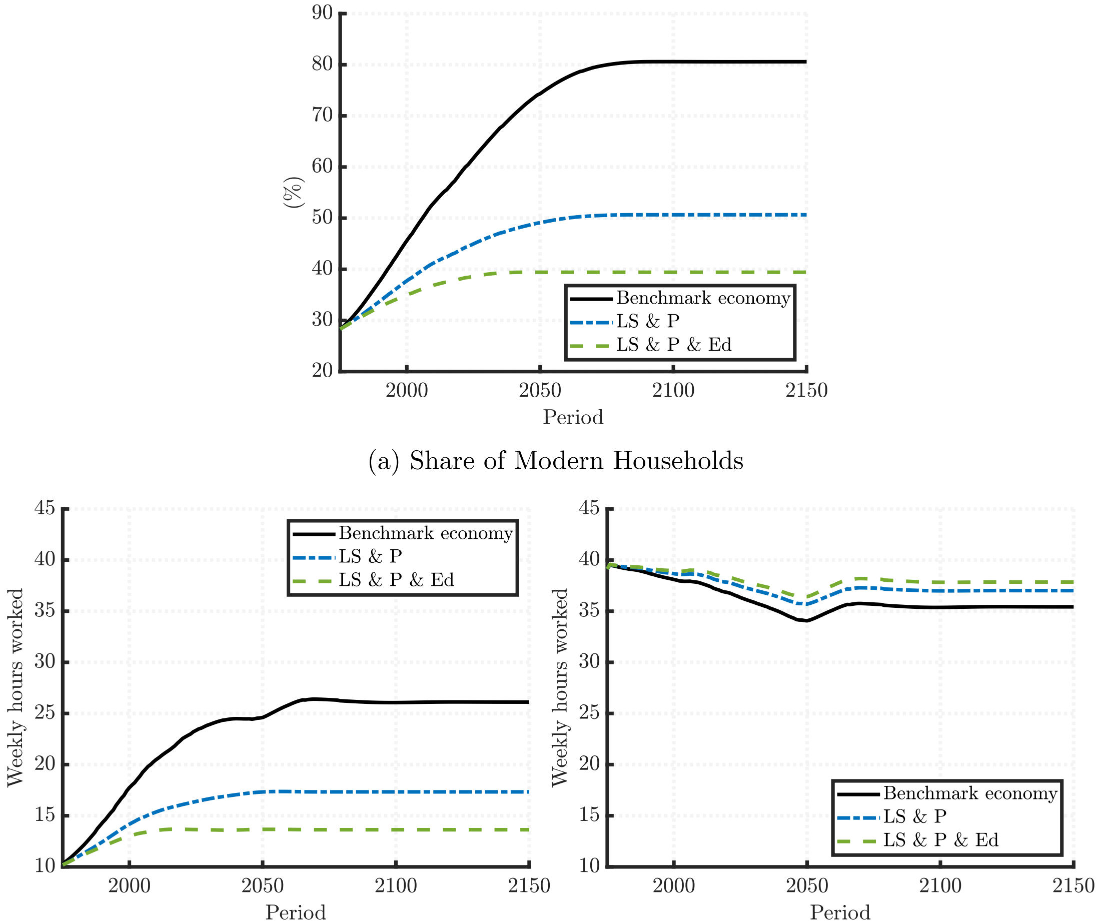
A3.5 Counterfactual Exercises

To better understand the role of education and women's employment separately, I simulate two economies differing in the policy functions of female labor supply and the educational transition. In the first model economy, which I label $LS\&P$, I assume that the policy functions for employment and hours worked by women are fixed to the ones at the initial steady state. In the second model economy, which I label $LS\&P\&Ed$, I remove the educational transition from the first model economy. I assume that the distribution of households by education is equal to the one in the initial steady state.

The consequences for labor supply. Figure A3.7 compares the transition paths of the share of modern households and hours worked by gender in the Benchmark, the $LS\&P$, and the $LS\&P\&Ed$ economies. Panel (a) shows that the share of modern households increases less in these two counterfactual economies compared to the Benchmark economy. It reaches 80%, 50.7%, and 39.5% in the final steady state for the Benchmark, $LS\&P$, and the $LS\&P\&Ed$ economy, respectively. Then, Panel (b) shows that, on average, women work less in these two economies compared to the Benchmark economy. In the final steady state, women work on average 26.1, 17.3, and 13.6 hours per week for the Benchmark, $LS\&P$, and the $LS\&P\&Ed$ economy, respectively. The rise in the share of modern households and hours worked compared to the initial steady state is attributable to women having fewer children during the transition. This reduces the opportunity cost of working since they incur lower time and monetary costs. Additionally, for the $LS\&P$ economy, the increase in the educational attainment of households also induces more women to participate and to work more. Finally, Panel (c) indicates that, on average, men increase their weekly hours worked to offset the reduced female labor supply compared to the Benchmark economy. This substitution effect accentuates by omitting the educational transition. On average, in the final steady state, men work 35.4, 37, and 37.9 hours per week in the Benchmark, $LS\&P$, and the $LS\&P\&Ed$ economy, respectively.

The consequences for Social Security. Figure A3.8 compares the evolution of the Social Security contribution rate required to balance the budget in the Benchmark economy with the other two hypothetical economies. Panel (a) in Figure A3.8 shows the contributions rate until 2019. This figure shows that the Social Security contributions rate necessary to balance the budget is higher under these two counterfactual economies. In particular, in 2019, the contributions rate is 13.3%, 14.2%, and 15.1% in the Benchmark, $LS\&P$, and the $LS\&P\&Ed$ economy, respectively. The reason behind this result is that until 2019 in these two counterfactual economies compared to the Benchmark economy, female Social Security contributions drop while the pension expenditure remains constant. Although men's contributions also increase, they do not compensate for the drop for women. Figure A3.9 shows the percentage difference in contributions (left panel) and pension expenditure (right panel) between each of these counterfactual economies and the Benchmark economy. In 2019, female contributions to Social Security decrease by 18% and 28% in the $LS\&P$, and the $LS\&P\&Ed$ economy, respectively. At the same time, men's Social Security contributions rise by 10% and 15% in the $LS\&P$ and the $LS\&P\&Ed$

Figure A3.7 Labor Market Outcomes in the Benchmark, *LS&P*, and *LS&P&Ed* Economies



(a) Share of Modern Households

(b) Female Weekly Hours Worked

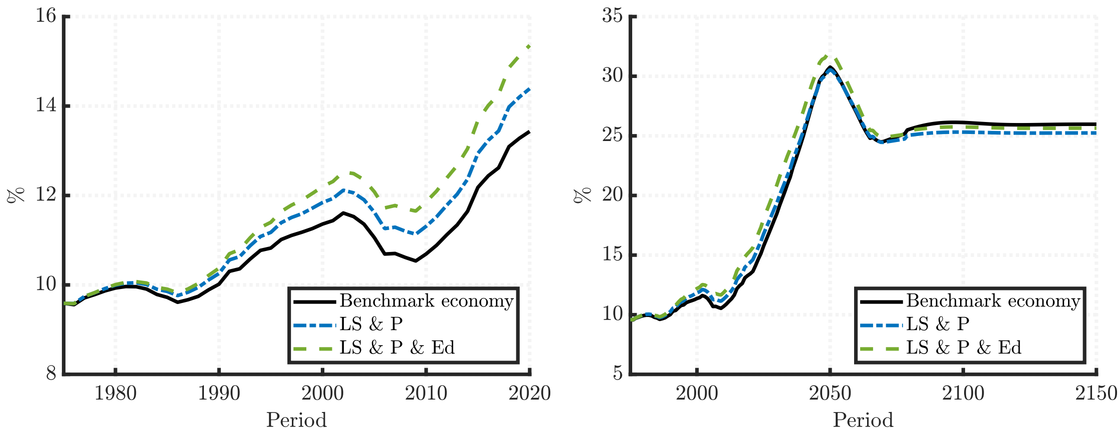
(c) Male Weekly Hours Worked

economy, respectively.

Panel (b) in Figure A3.8 shows a similar trend over the transition. However, in the final steady state, the tax rate is 25.94%, 25.64%, and 25.24% in the Benchmark, *LS&P*, and the *LS&P&Ed* economy, respectively. This is because female contributions to Social Security fall by more than the expenditure on female pensions. In particular, in the final steady state, compared to the Benchmark economy, female pension expenditure fall by 26.7% and 42.5% in the *LS&P*, and the *LS&P&Ed* economy, respectively while these drops for female Social Security contributions are 28.5% and 45.4%, respectively, see Figure A3.9. The reasons for this are that in these two economies, women are entitled to a higher share of minimum pensions and still receive widow pensions.

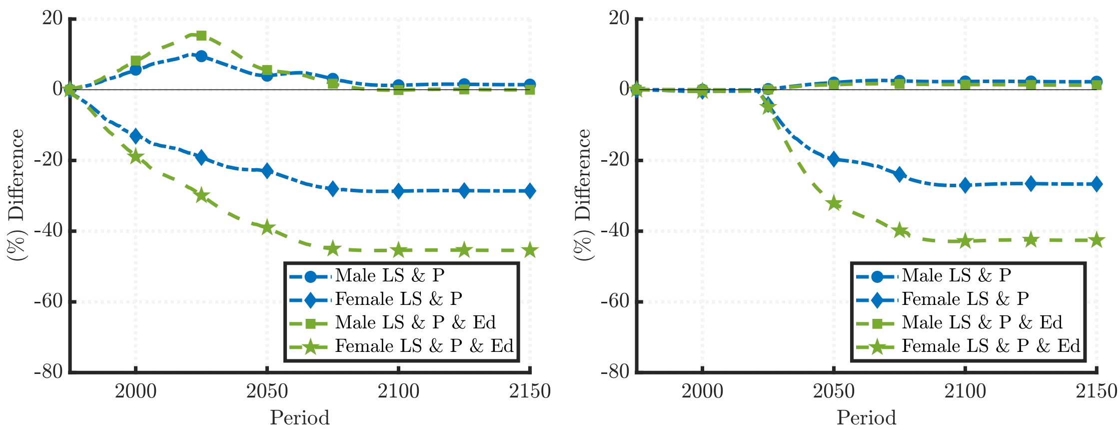
The consequences for the Gender Imbalance Indicator. Figure A3.10 compares the evolution of the FGII (left panel) and the MGII (right panel). By 2019 the drop in female labor force participation and hours worked compared to the Bench-

Figure A3.8 Social Security Contributions Rate in the *LS&P* and *LS&P&Ed* Economies



(a) Until 2019 (b) All the Transition

Figure A3.9 Percentage Change of Contributions and Pensions in the Counterfactual Economies Relative to the Benchmark Economy

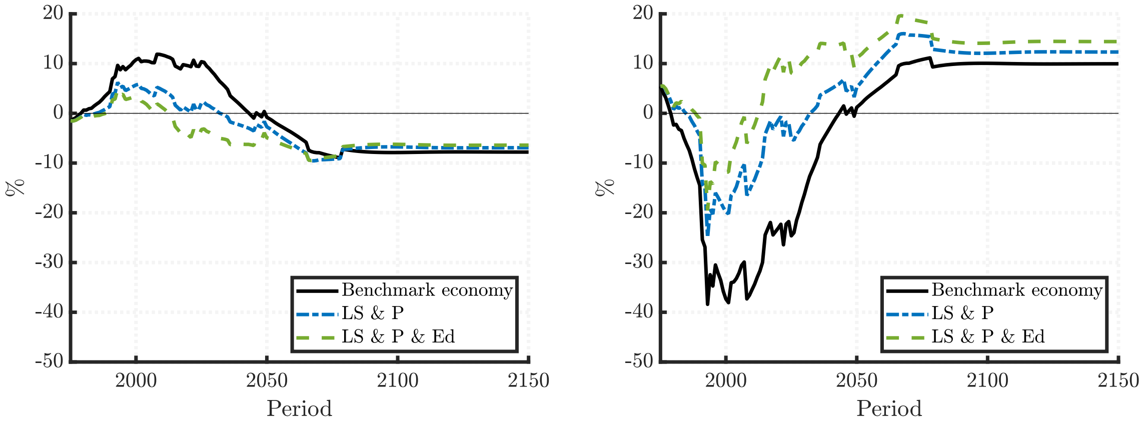


(a) Contributions (b) Pensions

mark economy explains all of the redistribution of resources from women to men. Moreover, the FGII falls to -4% if the educational transition is not included, indicating that 9.5% of women's pensions are financed by men in 2019. The role of women in pension funding is negligible in these two scenarios. In the final steady state, men finance a larger share of women's pensions, especially in the most restrictive economy without an educational transition. In other words, the educational transition amplifies the effects of a higher female labor force participation rate.

Overall, the Social Security System benefits from a higher female employment rate and educational attainment until 2019, preventing the payroll tax from increasing by 1-2 percentage points. This is still true until 2045. However, this is no longer the case in the long run since a lower tax rate is possible if female employment and educational attainment stay at the same level as the one in the initial steady state. The main reason is the drop in female pension entitlements and average female pen-

Figure A3.10 Gender Imbalance Indicator in the *LS&P* and *LS&P&Ed* Economies



(a) Female GII (b) Male GII

Note: The Female GII is the ratio of female contributions minus female pensions to male pensions. The Male GII is calculated as the male surplus over female pensions.

sions compared to the Benchmark economy. These hypothetical scenarios also show that men start redistributing resources to women earlier and at a greater scale than the Benchmark Economy.

Chapter 4

When Wives Command: Household Portfolio Choices and Marital Property Regime

4.1 Introduction

The marital property regime has been a key determinant of the economic nature of marriage. The degree of shared ownership of assets acquired during the marriage defines two broad types of marital property regimes: separate and community property. In separate property, each spouse maintains sole ownership of assets accumulated during the marriage and takes them upon dissolution. Contrary, in community property, most assets acquired during the marriage become jointly owned and split between spouses if the marriage ends.[1]

The type of marital property regime has relevant implications for savings decisions mainly because of two reasons. First, the marital property regime affects married couples' incentives to save because property division rules determine the allocation of spouses' savings *ex-post* marriage (Voena, 2015). While separate property limits the ability to tap into the spouse's savings, community property regulates that the common pool of assets accumulated during marriage must be shared in case of divorce, irrespective of who contributed the most to its acquisition. The different property division rules distort spouses' optimal savings decisions during the marriage, as spouses can differ in their contribution to household income or consumption levels. Second, property division rules also affect the economic cost of terminating the marriage (Imre, 2022). Unlike separate property, community property entails

[1] In Spain, under community property, labor income and profits earned by either spouse belong to the pool of commonly owned assets, while inheritance, gifts, and assets bought before marriage remain separate property. We denote this regime as community property or joint ownership throughout the paper.

a mandatory dissolution process involving an inventory of the common net assets, which is costly in terms of time and money.

An aspect that has received less attention in the literature is how property division rules interact with couples' financial portfolio choices. This paper fills this gap by investigating the impact of property division rules on household financial investment. The Spanish institutional setting serves as an ideal testing ground to address this question as the marital property regime law is regulated at the regional level, resulting in variation in the default rules across the Spanish regions. Separate property is the default regime in Catalonia and the Balearic Islands, while some form of community property is the default in the rest of the regions. Couples adopt the default marital property regime in their region of residence unless spouses agree on a different one by signing a prenuptial agreement. By means of an instrumental-variable (IV) strategy, we exploit this regional variation in marital law in combination with rich survey data from the Spanish Survey of Household Finances to provide causal estimates of the effects of property division rules on couples' financial portfolio choices. The Spanish Survey of Household Finances (or EFF for its acronym in Spanish) provides information on Spanish households' wealth, debt, and demographics. Particularly relevant for our study, it contains detailed information on household financial investment by asset class (i.e., bank deposits, shares, bonds, etc.) and on the marital property regime when households consist of married couples.

We find that separate-property couples take significantly more financial risk when wives are most knowledgeable about household finances. In particular, we find that separate property couples are 9% more likely to participate in risky assets than their counterparts married under community property when wives are the household heads. The definition of the household head in the EFF makes it very likely that this household member is the primary decision-maker regarding the household economy and finances. Specifically, the household head is the spouse most knowledgeable about the household economy and investments, being able to give detailed information about household wealth and debt holdings. We also find that separate-property couples hold more diversified portfolios towards risky assets than those married in community property. On average, couples married under separate property hold a share in risky asset classes 5 percentage points higher than couples married under community property when wives take a primary role in household finance investments.

Our identification strategy relies on assuming that the marital property regime affects financial outcomes only through the induced variation resulting from couples adopting the default regime in their region. However, the regional variation in default property regimes in Spain emanates from old legal traditions: Catalonia and the Balearic Islands adopted separate property during the Roman Empire's rule, while the other Spanish regions acquired community property from the Visigothic

Kingdom law system. Therefore, it is not unreasonable to think that the same legal traditions might have shaped attitudes towards risk or cultural norms differently between the two groups of regions over the course of history. We ensure that our results are robust to controlling for idiosyncratic differences that can affect household financial behavior and could have been captured by our instrument. In addition to including a wide range of socioeconomic characteristics and gaps between spouses, we show that our empirical findings remain stable and strongly significant when controlling for differences in risk aversion, financial sophistication, or gender norms promoting female financial independence.

To rationalize the empirical findings, we develop a two-period model of financial portfolio choice where couples differ in their marital property regime. For simplification, households consist of two spouses who are born married and face an exogenous probability of divorce. The household head decides on the level of consumption, which is public within the household, and her savings in safe *and* risky financial assets given her spouse's savings decisions and expectations about both spouses' future labor income, asset returns, and marital status.[2] In the model, property division rules dictate the asset allocation upon divorce and the corresponding dissolution costs. When separate property couples divorce, spouses take their individual assets according to the title of ownership and face no dissolution cost. In contrast, community property couples must incur dissolution costs as total household savings need to be equally split between spouses. We introduce this dissolution cost assuming that an exogenous fraction of total household income is destroyed in the event of divorce (Cubeddu and Ríos-Rull, 2003; Bacher, 2021b). Divorce represents a source of financial risk in the model because it requires couples to split their assets and because it results in a state with lower income levels and higher income risk. However, the strength of the precautionary savings motive differs across marital property regimes.

We calibrate the model to match key moments of Spanish married couples' financial behavior for which wives are the most knowledgeable about household finances. In particular, we calibrate the model assuming that wives are the ones making portfolio choices given their husband's savings decisions. By means of counterfactual simulations, we show that divorce risk and gender heterogeneity in labor income profiles are the most important determinants through which marital property regime affects financial portfolio choices. The model mimics the estimated marital property gap in risky financial investment in female-headed households. The model matches well both the targeted gap in participation in risky assets between marital property regimes and the untargeted gaps in the risky assets share and the total savings-to-income ratio. Relative to separate property, community property's higher marriage

[2] Our theoretical framework could be considered as a reduced-form version of the dynamic collective model of intra-household decision making (Mazzocco, 2005; Chiappori et al., 2002; Voena, 2015) where couples solve a constrained Pareto problem.

dissolution costs induce spouses to increase precautionary savings and lower their demand for risky assets. Low labor income levels and higher income risk for wives further strengthen couples' precautionary savings motive under divorce risk.

In the context of rising divorce rates in many countries, the data shows that women are exposed to greater labor income volatility and continue to accumulate less financial wealth than men, especially in risky assets (see, e.g., Global Gender Gap Report, 2022). Our research implies that a marital property regime that encourages an individual management of investment portfolios may be beneficial for wives. A diversification of the investment portfolio that encourages higher participation and share of risky assets allows insuring against the risk of divorce and the unpredictability of labor income dynamics with fewer savings.

Related literature and contributions.— A limited but growing literature has explicitly studied the implications of different marital property regimes for various household economic outcomes. Brassiolo (2013), Piazzalunga (2016), Imre (2022) and Huang et al. (2021) examine empirically how divorce laws interact with different marital property regimes in shaping households economic behavior. Like us, Imre (2022) exploit the regional variation in default marital property regime law in Spain. She investigates the effects of the marital property regime on female labor supply, fertility, marriage, and marital dissolution rates. We contribute to this literature by studying how property division rules shape household financial decisions.

This paper broadly complements the theoretical literature studying the interaction of marital transition dynamics and household savings behavior (see Yamaguchi et al. (2014); Voena (2015); Cubeddu and Ríos-Rull (2003); De Nardi et al. (2021)). Our paper is closely related to Voena (2015), who studies the interaction between property division rules and divorce laws in the US through the lens of a dynamic collective model of intra-household decision-making. Exploiting panel variation in U.S. divorce and property division laws, she finds that the parameter estimates of the model are consistent with a collective model where wives' share of household resources in marriage is low. This implies that women benefit from the laws that impose an equal division of property upon divorce, which gives community-property couples incentives to increase total asset accumulation and reduce wives' labor supply compared to separate property. Differently from Voena (2015), our theoretical framework nests into the class of unitary models of household decision-making but explicitly models how property division rules shape couples' financial portfolio allocation between safe and risky assets in the presence of uninsurable divorce and income risk. In this respect, we contribute to the literature studying how marital dynamics affect household portfolio allocation. Love (2010), Hubener et al. (2016) and Bacher (2021b) develop a joint framework of household structure and financial portfolio choice to study how couples and singles make portfolio choices following family shocks such as divorce or/and marriage. Our contribution here relies on introducing two types of property division rules in a theoretical portfolio choice framework and studying their implications for married couples' risky financial investments.

Our paper also contributes to the growing economic literature on gender and finance. In this literature, there is consensus regarding the fact that men invest more and less conservatively in financial assets than women because of differences in risk aversion (Bajtelsmit and Bernasek, 1996; Croson and Gneezy, 2009; Dohmen et al., 2011), financial literacy (Van Rooij et al., 2011; Lusardi and Mitchell, 2014; Hospido et al., 2021) or self-confidence (Barber and Odean, 2001; Bucher-Koenen et al., 2017; Klapper and Lusardi, 2020). More recently, the role of traditional gender norms has also been highlighted as another potential driver behind the gender gap in financial investment (Ke, 2021). Guiso and Zaccaria (2021) also show that more egalitarian norms increase household participation in financial markets, equity holdings, and asset diversification in Italy. Instead, we examine the impact of the marital property regime on household financial investment decisions, given the gender differences found in the previous literature regarding psychological traits, risk-taking, or social norms.

The rest of the paper proceeds as follows. The next section covers the Spanish institutional background. Section 4.3 presents the data, while section 4.4 empirically examines the role of the marital property regime for household financial behavior. Next, sections 4.5-4.9 lay down the theoretical model that rationalizes the empirical results. Section 4.10 offers concluding remarks.

4.2 Institutional Background

Spanish regions have considerable legislative autonomy. Particularly relevant for this paper, marital property regimes are regulated at the regional level. The marital property regime defines the legal ownership structure of assets acquired during the marriage. It regulates the division rule over couples' property upon marriage dissolution (due to divorce or death). Figure 4.1 shows that two marital property regimes coexist in Spain. While Catalonia and the Balearic Islands have separate property as their default property regime, some form of community property applies in the rest of the regions.[3] Under community property, assets acquired during the marriage are jointly owned, and they are split equally between the spouses upon marriage dissolution. By contrast, under separate property, each spouse retains full ownership of the assets they have acquired during the marriage in case of divorce or death.

The default marital property regime applies unless spouses agree on a different division rule signing a prenuptial agreement (*Capitulación Matrimonial* in Spanish).

[3] The Valencian Community, as an exceptional case, changed its default regime from community to separate property during the period 2008-2016.

Figure 4.1 Default Marital Property Regimes in Spain

Note: The figure plots the regional variation in default property regime across Spanish regions. Separate-property regions are Catalonia, and the Balearic Islands are in blue, while community-property regions are in green. Valencian Community changed to default separate property between 2008 and 2016.

Prenuptial contracts can be signed ex-ante or ex-post marriage, can be modified at any time during the marriage if both spouses agree and their monetary cost is relatively small (about 60 euros in 2021). Despite the simplicity of the procedure, most marriages merely adopt the default property regime in their region. Appendix Figure A4.1 shows the evolution of total prenuptial agreements as a share of marriages and prenuptial agreements for separate property as a share of total contracts in Spain. The number of prenuptial agreements remains below 20% of marriages. Among those prenuptial agreements, more than 90% corresponds to a change from community property to a separate property regime.[4] Figure A4.2 shows that both marital property regimes have similar marriage and divorce dynamics.

Community and separate property imply different costs of distributing marital assets between spouses ex-post marriage (i.e., divorce or death) (Imre, 2022). Unlike couples married under separate property, community-property spouses are required to dissolve the community property regime by law. The procedure requires making an inventory and valuing all common assets and liabilities, which requires both spouses' approval. Then, the ownership of half the net value of the shared pool of assets can be assigned to each spouse.[5] Therefore, divorce is more costly and lengthier for couples married under community property compared to those married under separate property.

[4] We find similar trends for the evolution of prenuptial contracts to adopt separate property by region.

[5] This procedure needs to be done before a public notary. The average cost ranged between 1,000 and 1,500 euros in 2022.

4.3 Data

We use household-level data from the Spanish Survey of Household Finances. The survey is conducted every two years by the Bank of Spain and spans from 2002 to 2020 (7 waves in total). The survey reports detailed information on households' income, wealth, portfolio composition, and a rich set of socio-economic characteristics based on personal interviews.

We exploit particular features of the EFF, which are rarely included in surveys reporting information about household wealth. First, the survey includes information on the marital property regime of couples, which is not available in other surveys such as the Bank of Italy's Survey of Households Income and Wealth (SHIW) or the Federal Reserve's Survey of US Consumer Finances (SCF). Second, the definition of the household head makes it very likely that he or she is the main decision-maker of the household economy and finances. The specific definition provided to households reads: "the person who knows more about the economy and finances of the household living at this address". Thus, the household head is the person who is the most knowledgeable about the household's finances, i.e. household income, expenditures, investments, assets, etc. It is not simply a household member, but who is in charge/knows the most about the household's finances.

We restrict the estimation sample to married couples over 25 years old with both spouses employed so that both contribute to household income. We drop self-employed workers because their financial decisions are most likely to be determined by other motives than the general population. For instance, self-employed individuals tend to opt for the separation of property because this regime provides a way of sheltering a fraction of household assets from the risk of bankruptcy.

Table 4.1 reports summary statistics of our sample. Panel A presents summary statistics of households' socioeconomic characteristics. About 75% of couples are married under community property. This is not surprising since all Spanish regions have community property as the default marital property regime except for two. In addition, wives take a more prominent role in managing household finances in about one-third of households, independently of the marital property regime. On average, the spouse most knowledgeable about the household finances (i.e. the household head) is 46 years old, more educated, slightly older, and earns more than his/her spouse. Looking at the differences in socioeconomic characteristics between the two types of regimes, we can observe that, on average, the household head in separate-property couples is more educated and more likely to work in the financial sector. In addition, these couples are wealthier and earn a higher income compared to their counterparts married under community property. Panel B presents summary statistics of household financial outcomes. We classify shares and mutual funds as risky financial assets, while fixed-income securities, savings, and checking accounts are categorized as safe financial assets. Panel B shows that separate property couples' average participation rate in risky assets and the risky portfolio share is higher.[6]

Appendix Tables A4.1 and A4.2 reproduce the summary statics by gender of the household head. The average differences in socioeconomic characteristics and financial outcomes hold irrespective of the gender of the household head except for the wage differential between spouses. Male household heads earn about twice as much as their spouses, while female household heads earn less. Notice that since the percentage of female household heads who are also second earners is 72%, compared to only 18% for men, it is virtually the same analyzing the differential behavior of second earners or women. Finally, it is worth noticing that the gap in risky investment is considerably larger for households led by females.

4.4 Empirical Results

To investigate whether property division rules in marriage affect couples' risky financial investment, we rely on an instrumental variable strategy. The choice of marital property regime is potentially endogenous, as spouses can opt out of the default regime by signing prenuptial contracts. Frémeaux and Leturcq (2020) show using French administrative data that separate property could be used strategically by the wealthiest spouse to protect their wealth in case of divorce in unequal partnerships. If wealthier couples self-select into separate property, regressing directly financial participation in risky assets on a separate property dummy would overestimate the effects of this property division rule on risky financial investment.

In our sample, 86% of households living in community-property regions adopted the default regime. This means that around 14% of couples in this group of regions changed to separate property. Figure A4.3 in the Appendix disaggregates the share of households opting out of community property by net wealth percentile and shows that couples in the highest percentile are more likely to choose separate property. To avoid this source of endogeneity in our setting, we exploit the regional variation in default regimes across Spanish regions and use the region of residence as an instrument for marital property regime as follows:

$$Y_{i,t} = \beta_0 + \beta_1 \text{Sep. Property}_{i,t} + \beta_2 \text{Female}_{i,t} + \beta_3 (\text{Sep. Property} \times \text{Female})_{i,t} + \delta' X_{i,t} + \lambda_t + \upsilon_{i,t} \quad (4.1)$$

$$\text{Sep. Property}_{i,t} = \alpha_0 + \alpha_1 \text{Region}_{i,t} + \gamma' X_{i,t} + \lambda_t + \varepsilon_{i,t} \quad (4.2)$$

where Sep. $\text{Property}_{i,t}$ equals 1 if household i is married under separate property and 0 if married under community property, while $\text{Region}_{i,t}$ equals 1 if the couple

[6]The high participation rates are driven by the fact that the EFF survey oversamples at the top of the wealth distribution

TABLE 4.1 Household Summary Statistics

	Mean	St. dev.	Separate	Community
Panel A. Socioeconomic characteristics				
Household head				
Separate property	0.26	0.44		
Female	0.34	0.47	0.32	0.35
Age	46	8.69	46	46
Education				
Less than high school	0.23	0.43	0.16	0.26
High School	0.34	0.47	0.31	0.35
College	0.43	0.49	0.53	0.39
Occupation in financial sector	0.05	0.22	0.08	0.04
Comparative ratios bw spouses				
Education ratio bw spouses	1.10	0.48	1.10	1.11
Age ratio bw spouses	1.03	0.10	1.04	1.03
Wage ratio bw spouses	1.58	1.82	1.74	1.53
Other controls				
Home-ownership				
Rent	0.09	0.29	0.10	0.09
Ownership	0.87	0.33	0.86	0.88
Other	0.04	0.18	0.05	0.03
Household size	3.52	0.99	3.47	3.53
Income (thousands eur)	66.95	92.96	90.35	58.79
Net wealth (thousands eur)	552.02	3418.54	1123.63	351.35
Panel B. Financial Variables				
Financial Variables				
Participation risky assets	0.30	0.48	0.38	0.27
Risky asset classes (%Total asset classes)	0.15	0.24	0.19	0.14
Risky assets share	0.15	0.29	0.21	0.13

Note: This table shows summary statistics for two-spouse households characteristics and by marital property regime of the household head. The sample includes information from the 2002-2020 waves of the Spanish Survey of Household Finances and is restricted to two-spouse households aged above 25 years old who are employed. Self-employed households are excluded from the sample. Observations: 4910 (4800 for the education ratio, 4791 for the risky asset classes share, and 4774 for the risky assets share)

lives in Catalonia or the Balearic Islands and 0 if otherwise. The main identifying assumption is that couples' region of residence is correlated with their marital property regime choice but uncorrelated with household financial portfolio choices. To investigate whether there are heterogeneous effects depending on the gender of the household head, we add an indicator variable, $\text{Female}_{i,t}$, that equals 1 if the household head is the wife and its interaction with the property division rule variable.

We additionally control for a full range of household socio-economic characteristics, X_{it}, including household income and net wealth deciles, number of individuals living in the household, household head's age, education, homeownership, civil union status, occupation in the financial sector and comparative proxies between spouses (education, age, and wage ratios). Finally, we include survey year λ_t fixed effects to capture time trends affecting household financial investment.

Table 4.2 reports the first-stage results. The coefficients are positive and statistically significant, suggesting that living in Catalonia or the Balearic Islands is strongly correlated with being married under separate property. This, together with the high F-stat values, confirm the relevance of our instrument. Table 4.3 presents the 2SLS estimation results. Consistent with the literature on gender differences in finance, the negative coefficients for the female dummy indicate that couples are less likely to take financial risks when wives take a more prominent role in managing household finances compared to husbands. However, property division rules introduce significant differences in the participation and portfolio diversification of risky assets among female-headed couples. In particular, households married under a separate property regime are 9% more likely to invest in risky assets than their community property counterparts when wives are the most knowledgeable about household finances. These couples also hold a share in risky asset classes up to 5 percentage points higher compared to couples married under community property.

Table 4.2 First-stage Regressions

	(1) Sep. Property	(2) Sep. Property × Female
Regions with Default Sep. Property	0.542***	
	(0.016)	
Regions with Default Sep. Property × Female		0.541***
		(0.029)
Household Characteristics	Yes	Yes
Survey FE	Yes	Yes
F-value	103.223	46.941
Prob > F	0.000	0.000
Observations	4262	4262
R^2	0.341	0.413

Note: The sample includes all two-earner married households in 2002-2020 except for households living in Valencian Community since this region changed the default marital property regime law between 2008-2016. This table provides results of the first-stage regression of the separate-property variable on a dummy variable that takes a value equal to 1 when the couple's region of residence is Catalonia or the Balearic Islands. Standard errors are robust.

Table 4.3 Instrumental Variables Estimates

	(1) Risky Financial Assets	(2) % # Risky Financial Asset Classes
	IV-2SLS	IV-2SLS
Separate Property	-0.059	-0.037
	(0.041)	(0.026)
Female	-0.075***	-0.039***
	(0.015)	(0.010)
Female × Sep. Property	0.090***	0.051***
	(0.022)	(0.015)
Households Characteristics	Yes	Yes
Survey Year FE	Yes	Yes
Observations	4262	4156

Note: The sample includes all two-earner married households in 2002-2020. This table provides 2SLS results from a model where the dependent variable is a binary variable that equals 1 if households hold wealth in risky assets (i.e., listed shares, unlisted shares, and mutual funds) (Column (1)) or the share of different risky asset classes (Column (2)). Separate property is instrumented using a dummy for residence in Catalonia or the Balearic Islands. *Female* is a dummy variable that equals 1 if the household headship is female and 0 otherwise. We exclude from the sample couples living in Valencian Community as this region changed its default regime during the time period considered. Standard errors (in parenthesis) are robust and clustered at the regional level.

4.4.1 Robustness Checks

In our context, the exclusion restriction implies that property division rules affect financial outcomes only through the induced variation resulting from couples adopting the default regime in their region of residence. The most relevant threat to identification in our setting is that regional variation in default regimes captures cultural differences that might affect household financial behavior beyond property division rules themselves. The multiple marital property regimes result from different legal traditions: Catalonia and the Balearic Islands adopted separate property during the Roman Empire's rule, while other Spanish regions acquired community property from the Visigothic Kingdom law system. It is not unreasonable to think that such old legal traditions have shaped local cultural patterns differently, and this could translate into different household financial behavior. We exploit the information provided in the EFF survey to control for some of these potential confounders.

Different legal traditions could have influenced preference towards risk and financial sophistication levels. They can also promote or discourage female financial independence, which can be transmitted through family ties from generation to generation. Imre (2022) provides evidence on this channel by showing that separation of property promotes a higher female labor supply in Spain. We use a variable measuring financial risk-taking as a proxy for household risk aversion, online banking and ownership of managed financial accounts as proxies for financial sophistication, and labor supply of household heads' mothers as a proxy for gender norms promoting female financial independence. Table A4.3 and A4.4 present 2SLS estimates when controlling for risk attitudes, financial sophistication levels, and egalitarian gender norms and show that our results are robust to these alternative channels.

4.5 Theoretical Framework

We develop a two-period unitary household financial portfolio choice model to shed light on the mechanisms behind our empirical findings. Households consist of two individuals, $i = \{h, w\}$, who live for two periods and are born married. In the first period, both spouses are subject to idiosyncratic labor income shocks. The household head decides household consumption, which is a public good, and her allocation of savings between a risk-free and a risky asset. For simplicity, the spouse's savings and portfolio choices are exogenous. In the second period, couples face an exogenous probability of divorce and idiosyncratic labor income shocks. The marital property regime only matters for the allocation of assets between spouses in case of divorce and the dissolution costs of marital assets. Under community property, the sum of the spouses' total assets is divided equally between them. Moreover, spouses have to pay a dissolution cost of marital assets. In contrast, separate property spouses keep the property of their individual assets and pay no dissolution cost as there is no common pool of assets to be divided.

4.5.1 Preferences

Households have a time-separable CRRA preference over consumption, c. The period flow utility is given by

$$u(c) = \frac{c^{(1-\gamma)}}{1-\gamma}, \tag{4.3}$$

where γ denotes the coefficient of relative risk aversion.

4.5.2 Asset Returns

The safe asset earns a constant gross return r_s, and the risky asset a random gross return r_r. We assume the return of the risky asset follows a normal distribution $r_r \sim N(\mu_r, \sigma_r^2)$, is independent and identically distributed and such that $\mu_r > r_s$.

4.5.3 Income Profiles

Income y^i for spouse i can be split into a deterministic and a stochastic component:

$$y^i = \bar{y}^i \epsilon^i, \tag{4.4}$$

where $\bar{y}^i$ represents the deterministic gender specific component and ϵ^i is the stochastic component. In particular, we assume that the stochastic component follows an AR(1) process:

$$\ln(\epsilon^{i'}) = \rho\epsilon^i + \upsilon; \quad \upsilon \sim \mathcal{N}\left(0, {\sigma_i}^2\right). \tag{4.5}$$

4.5.4 Divorce and Marital Property Regime

In the second period, couples face an exogenous divorce probability, δ. The probability of divorce is common across marital property regimes. If couples divorce, the allocation of marital assets between spouses and the corresponding dissolution costs depend on the marital property regime, m. When couples are married under community property, $m = c$, they split total assets equally and have to pay a dissolution cost of marital assets, κ^i. This cost accounts for time and all legal fees spouses must pay to the public notary to dissolve the shared pool of marital assets (i.e., inventory, valuing the assets, etc.). In contrast, couples married under separate property, $m = s$, take their individual assets upon divorce and pay no dissolution costs.

4.5.5 Timing

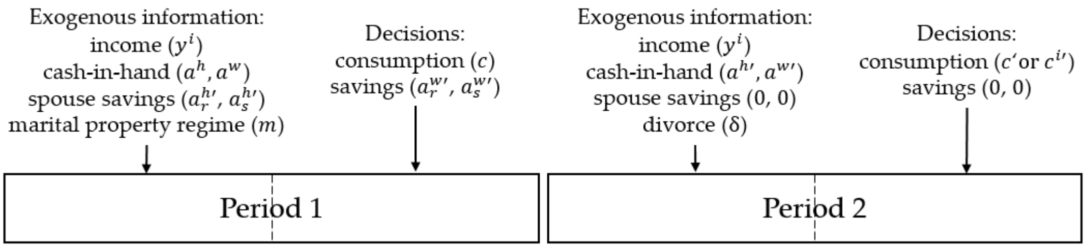
Figure 4.2 shows a timeline with the sequence of events in the model. In the first period, the household head learns both spouses' current income realization, her spouse's savings decisions, and marital property regime. Afterward, she decides on consumption, which is public within the household, and her allocation of savings between safe and risky assets. In the second period, the household head learns the spouse's income realizations, the spouse's cash-on-hand, and whether the couple divorces. Then, she decides optimally to consume all available resources.

4.5.6 Recursive Formulation

Notice that as the risky asset follows an i.i.d process, we can combine safe and risky assets into one "asset cash-in-hand" state variable: $a = (1 + r_r)a_r + (1 + r_s)a_s$

The state variables for a couple are the household head's asset cash-on-hand (a^i), her spouse's asset cash-on-hand (a^j), her spouse's choices of risky and safe assets

Figure 4.2 Timing of Events in the Model



Note: For simplicity we do not incorporate the state variables marital property regime in the second period-as it is constant over the life cycle- and the divorce in the first period -since households are always married in this stage.

$(a_s^{j'}, a_r^{j'})$, both stochastic components of income realizations (ϵ^i, ϵ^j) and their marital property regime (m).

The corresponding value function of married couples is as follows:

$$V^M\left(a^w, a^h, a_s^{h'}, a_r^{h'}, \epsilon^w, \epsilon^h, m\right) = \max_{a_s^{w'}, a_r^{w'}, c} \frac{c^{(1-\gamma)}}{1-\gamma}$$
$$+\beta\left[(1-\delta)\,\mathbb{E}V^M\left(a^{w'}, a^{h'}, 0, 0, \epsilon^{w'}, \epsilon^{h'}, m\right) + \delta \sum_{i=w,h} \mathbb{E}V^D\left(i, a^{w'}, a^{h'}, 0, 0, \epsilon^{i'}, m\right)\right]$$

$$\begin{aligned}
& c + \sum_{i=w,h} a_s^{i'} + \sum_{i=w,h} a_r^{i'} = \sum_{i=w,h} y_t^i + \sum_{i=w,h} a^i \\
& a^{i'} = (1+r_r)\, a_r^{i'} + (1+r_s)\, a_s^{i'}, \quad \forall i = \{w, h\} \\
& y^i = \bar{y}^i \epsilon^i, \quad \forall i = \{w, h\} \\
& \epsilon^i \sim \mathcal{N}\left(0, {\sigma_i}^2\right), \quad \forall i = \{w, h\} \\
& r_r \sim N(\mu_r, \sigma_r^2) \\
& \mu_r > r_s \\
& \epsilon^i \perp r_r, \quad \forall i = \{w, h\}
\end{aligned} \tag{4.6}$$

Similarly, the value function of a divorced individual i in the second period is:

$$\begin{aligned}
& V^D\left(i, a^{w'}, a^{h'}, 0, 0, \epsilon^{i'}, m\right) = \max_{c^{i'}} \frac{\left(c^{i'}\right)^{(1-\gamma)}}{1-\gamma} \\
& c^{i'} = \begin{cases} y^{i'} + \frac{a^{w'}+a^{h'}}{2} - \kappa^i & \text{if } m = c \\ y^{i'} + a^{i'} & \text{if } m = s \end{cases} \\
& y^{i'} = \bar{y}^i \epsilon^{i'} \quad \epsilon^{i'} \sim \mathcal{N}\left(0, {\sigma_i}^2\right)
\end{aligned} \tag{4.7}$$

4.6 Calibration

We calibrate the model using a two-step strategy. In the first step, we use data to estimate the parameters that can be identified outside the model. In the second step, we calibrate the remaining parameters to match the empirical participation gap in risky assets between separate and community-property couples. In the baseline calibration, women are assumed to be the household head. Table 4.4 summarizes the main parameter values.

Table 4.4 Parameters Calibrated

Parameter	Value	Source
First step		
$\bar{y}^h$	23958.72	EFF
$\bar{y}^w$	19166.88	EFF
σ_h^2	0.541	EFF
σ_w^2	0.609	EFF
ρ_h	0.571	EFF
ρ_w	0.531	EFF
σ_r^2	0.206	Bank of Spain
μ_r	0.0203	Bank of Spain
r_s	0	See text
δ	0.24	INE
γ	10	Cocco et al. (2005)
β	1	See text
α_1	24.12%	EFF
α_2	7.60%	EFF
Second step		
κ	10%	-

Starting with the first-step parameters, we set the permanent component of income $\bar{y}^i$ to match the average gender wage gap between spouses observed in the EFF data between 2002 and 2020. We focus on working married couples for which wives are the most knowledgeable about household finances (i.e., female-headed households), which gives us a gender wage gap of $\frac{\bar{y}^h}{\bar{y}^w} = 1.25$. Regarding the stochastic component of the income process, we estimate the following regression using the panel structure of the EFF:

$$\ln w_{jt}^i = \beta_1 age_{jt}^i + \beta_2 (age^2)_{jt}^i + \beta_3 occupation_{jt}^i + \lambda_j + u_{jt}^i \quad \forall i \in \{h, w\} \tag{4.8}$$

where w_{jt}^i denotes the monthly wage of spouse i in household j and λ_j refers to household fixed effects. We then regress the residuals obtained from this estimation on their time lags to obtain the persistence parameters of the AR(1) process for the stochastic shocks and the variance of the innovations. Table 4.4 presents the estimates of these two objects. The estimates indicate that married women's labor income is more volatile than their husbands'. Females' labor income variance is higher, and the persistence of their stochastic income process is somewhat lower. When solving the model numerically, we discretize the labor income shock using the Tauchen (1986) method.[7]

The average return of the risky asset takes the value $\mu_r = 2.03\%$, and its variance $\sigma_r^2 = 0.206^2$, consistent with average annual total returns and volatility of the IBEX-35 index between 2002-2021.[8] For simplicity, we set the net return of the safe asset to 0, $r_s = 0$.

The divorce probability is set to 24%, a linear interpolation between the average divorce rate for marriages over 5 years old (18%) and the maximum divorce rate of 30% for marriages over 20 years old. The interpolation brings the probability of divorce closer to that observed for couples married for more than 15 years, which reflects that in our sample the average age of first marriage is 31 while the average age is 47 years old. We used the Divorce Indicators data starting in 2005 from the Spanish Statistics National Institute (INE for its acronym in Spanish).

We borrow the risk aversion parameter from Cocco et al. (2005) and set it to $\gamma = 10$. Regarding the discount factor, we set $\beta = 1$ as our theoretical model has only two periods.[9]

The last first-step calibrated parameters are husband savings. The data from the EFF survey only provides information on household-level wealth holdings rather than individual savings. Since savings patterns and portfolio choices differ between married and single individuals (Bacher, 2021a; Love, 2010; Bertocchi et al., 2011), it would be misleading to use the data for single individuals to calibrate married men's savings profiles. To overcome this challenge, we assume that the contribution of each spouse to household savings is proportional to their labor income. This implies that the distribution of savings between spouses is proportional to their wage gap.[10] Formally, let's denote α_1 husband's total savings. We compute this share as follows:

$$\frac{a^h}{y^w + y^h} = \alpha_1 \times \frac{a}{y^w + y^h} \quad \text{where } \alpha_1 \equiv \frac{1}{1 + \frac{y_w}{y_h}} \tag{4.9}$$

where $\frac{a}{y} \equiv \frac{a}{y^w+y^h}$ is retrieved from the panel structure of EFF data 2002-2020 for

[7] In particular, we discretized the income shock using ten grid points.

[8] Series 'Cotización y contratación. Acciones. Sociedad de Bolsas y Sociedad Rectora de la Bolsa de Madrid. Índice cotización. Indice IBEX 35' downloaded from www.bde.es.

[9] See Gomes et al. (2021) for a literature discussion of the estimates of the coefficient of risk aversion, discount factor, and participation costs in asset allocation models over the life cycle.

[10] Grabka et al. (2015) and Meriküll et al. (2021) show using German and Austrian individual-level data that labor earnings are one of the main factors explaining spouses' share in total household savings.

households with finances led by wives. Specifically, we use the average change in total household financial savings between two consecutive waves to measure a while income refers to annual labor earnings. We obtain $\alpha_1 = 24.12\%$. Appendix Table A4.6 compares these shares with the total household savings to income ratio. Additionally, we use the portfolio share in risky assets of divorced men in the sample to calibrate that of husbands'. This implies a risky portfolio share for husbands of $\alpha_2 = 7.60\%$.

In the second step, we use the one remaining parameter, i.e., the dissolution cost of marriage κ^i, to target the gap in risky asset participation between households married under separate and community property regimes. Recall that the dissolution cost of marriage is only paid by community property couples. We introduce the individual cost in the model as follows:

$$\kappa^i = \kappa y^{i'} \tag{4.10}$$

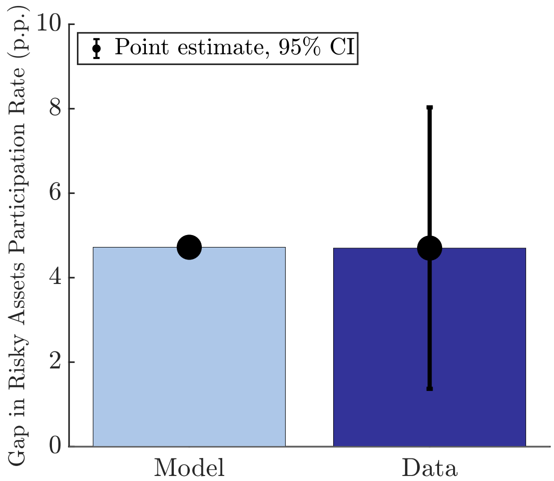
where y^i refers to the labor income of spouse i and κ, represents the fraction of total income destroyed in the event of marital dissolution. We set $\kappa = 10\%$, which falls below the range of values explored in previous studies such as Cubeddu and Ríos-Rull (2003) for the US economy.

4.7 Model Results

We begin quantitatively assessing the match of the model to the data. Figure 4.3 compares the gap in participation in risky assets between marital property regimes generated by the model and estimated in the data for couples whose household finances are led by wives. The model matches the targeted moment very well: it predicts a participation gap in risky assets between separate-property and community-property couples of 4.7 percentage points (pp) which equals exactly the estimated gap in the data. The empirical counterpart is estimated regressing female-headed households' participation in risky assets on a separate property regime dummy. To be consistent with our empirical strategy described in Section 4.4, we instrument the property regime variable with households' region of residence in Catalonia and Balearic Islands and control for the full range of socio-economic characteristics. Column (2) in Appendix Table A4.5 shows the results of this estimation.

Figure 4.4 presents the model fit for the gap in risky assets share and total savings-to-income ratio between the different marital property regimes. Notice that these gaps are untargeted in the calibration exercise. Again, the empirical counterparts are based on regressions for female-headed household outcomes on an indicator variable representing the marital property regime. Columns (2) and (3) in Appendix Table A4.5 show the 2SLS estimates of these two savings outcome gaps, respectively. The simulated model outcomes slightly underpredict the positive gap in the share

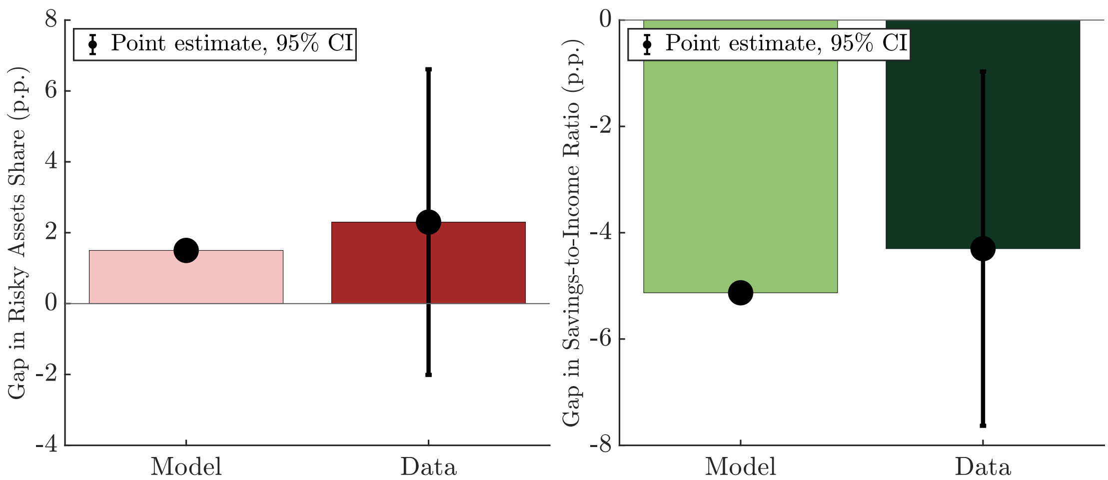
Figure 4.3 Property Regime Gap in Participation in Risky Assets: Model vs. Data



Note: This figure plots the property regime gap in the participation in risky assets generated by the model and the one estimated in the data. The gap is computed as the difference between separate-property and community-property households' portfolio share in risky assets. The darker blue bar refers to the 2SLS estimate of the gap and the corresponding 95% CI using EFF survey waves 2002-2020. The lighter blue bar refers to the model simulation outcome.

of risky assets (Figure 4.4a) and slightly overpredicts the negative gap in financial savings (Figure 4.4b). Nonetheless, the model results fall within the 95% confidence interval.

Figure 4.4 Property Regime Gap in Risky Assets Shares and Total Savings: Model vs. Data



(a) Risky Assets Share (b) Total savings-to-income Ratio

Note: This figure plots the property regime gap in the portfolio share in risky financial assets and the total savings-to-income ratio generated by the model, and the one estimated in the data. The gap is computed as the difference between separate-property and community-property households' outcomes. The darker bar refers to the 2SLS estimate of the gap and the corresponding 95% CI using EFF survey waves 2002-2020. The lighter bar refers to the model simulation outcome.

4.8 Explaining the Gap in Risky Investment

4.8.1 Transmission Channels

We now study the channels through which the marital property regime affects households' investment choices by means of counterfactual simulations. To do so, we change the parameter values of interest, solve the model again, and contrast the resulting simulation outcome to the baseline economy.

Divorce is a key driver of the marital property regime gaps in the model as property division rules dictate the sharing rule of assets between spouses upon divorce as well as the dissolution costs of marriage. Without divorce risk, couples face the same optimization problem during marriage, and their optimal portfolio choice decisions should be the same. Table A4.9 in the Appendix shows that risky asset share, participation rate, and total savings gap collapse to 0 when shutting down the divorce risk (i.e., $\delta = 0$).

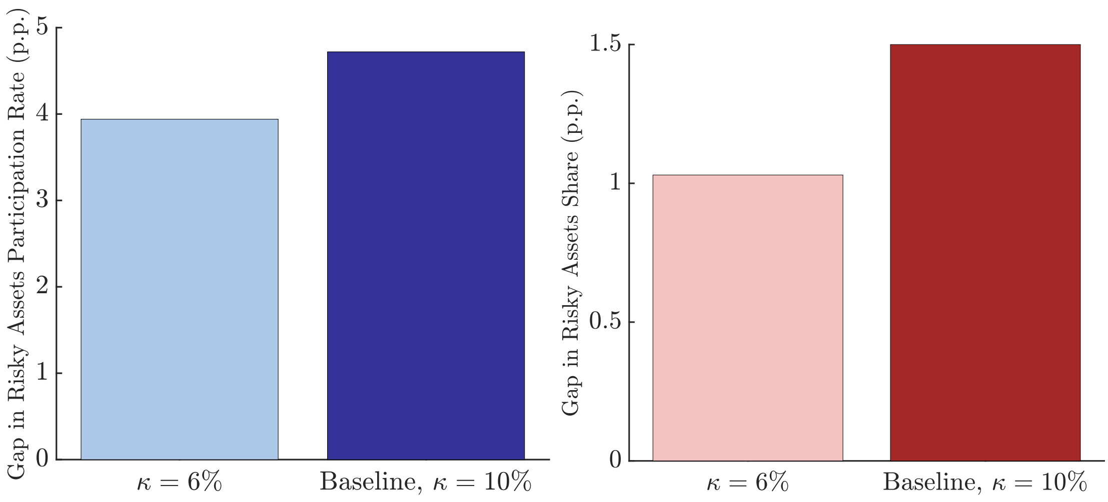
The dissolution costs of marriage are a source of heterogeneity across marital property regimes. In the model, we assume that community-property couples must pay the cost of dissolving the common pool of assets while separate-property couples do not. The strength of the precautionary savings motive increases with the dissolution costs of marriage (i.e., the proportion of permanent income destroyed in the event of divorce). Figure 4.5 shows the model simulation outcome for the gap in the risky assets participation rate and the risky share for a lower value of κ. As can be inspected, both gaps increase with the dissolution costs of marriage as wives married in community property demand more safe assets to self-insure against divorce risk.

To explore how income level differences between spouses affect the property regime gaps in risky financial investments, we simulate a counterfactual scenario where we invert the gender income gap in permanent income (i.e. $\frac{\bar{y}^h}{\bar{y}^w} = 0.80$).[11] Figure 4.6 shows that the gap in risky investment both at the extensive and intensive margin decreases as the wife's permanent income increases relative to their husband's. In fact, it becomes slightly negative. Notice that in this alternative economy, all married women would experience a smaller drop in consumption in case of divorce compared to the baseline economy as they earn higher permanent income on average. Thus, divorce becomes less risky for those married under community property, which reduces their demand for safe assets.

Finally, we investigate how income risk shapes the marital property regime gap in risky financial investments. We do so by assigning wives the stochastic part of their husbands' labor income process (variance and persistence), lowering their expo-

[11] For coherence, we also change the calibration for the husband's savings as we assume that spouses' distribution of household savings during the marriage is proportional to the wage gap.

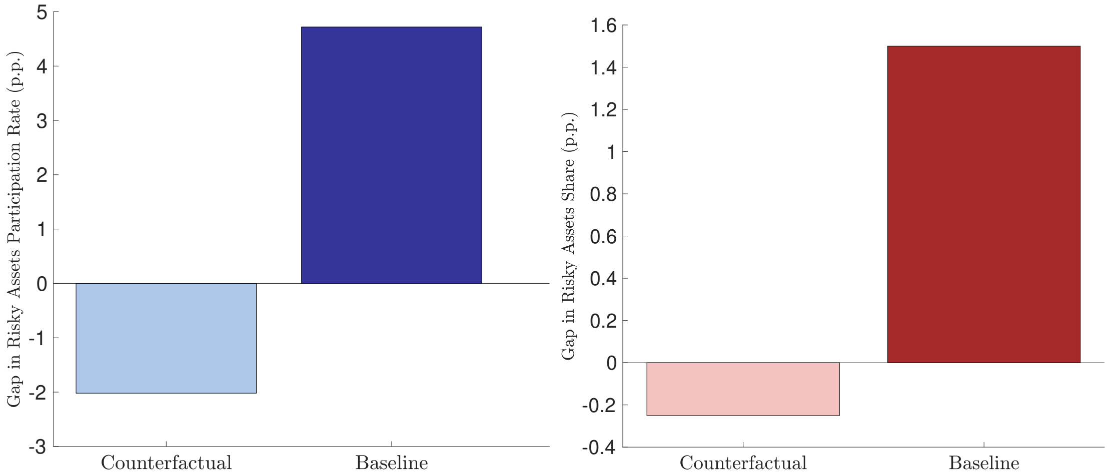
Figure 4.5 Counterfactual Scenario: Alternative Dissolution Costs of Marriage



(a) Risky Assets Participation Rate (b) Risky Asset Share

Note: This figure plots the property regime gap in the participation and portfolio share in risky financial assets in the counterfactual scenario and the baseline economy. The gap is computed as the difference between separate-property and community-property households' outcomes.

Figure 4.6 Counterfactual Scenario: Alternative Income Levels

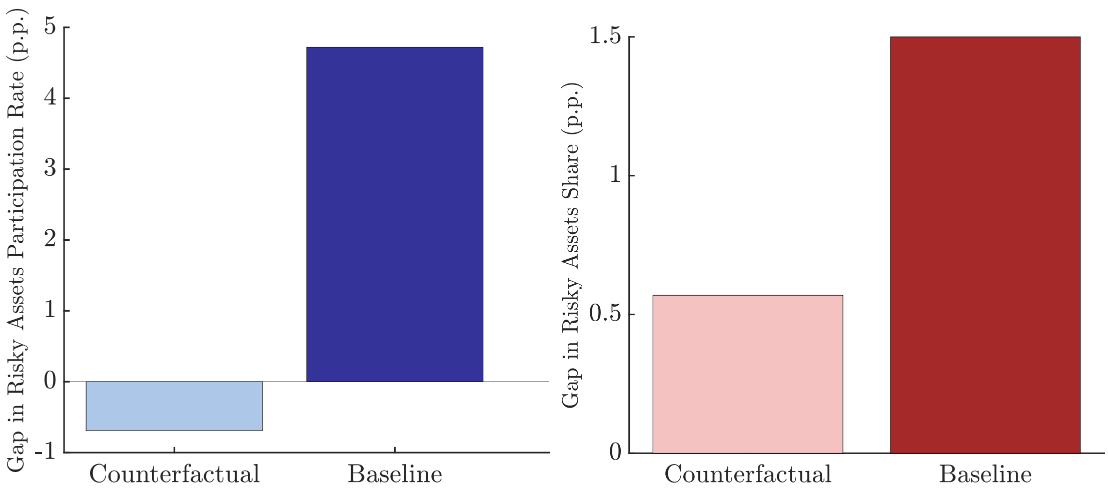


(a) Risky Assets Participation Rate (b) Risky Asset Share

Note: This figure plots the property regime gap in the participation and portfolio share in risky financial assets in the counterfactual scenario and the baseline economy. The gap is computed as the difference between separate-property and community-property households' outcomes.

sure to income fluctuations. Figure 4.7 shows that the gap in risky investment gets significantly reduced both at the extensive and intensive margin. Even becoming slightly negative for the participation rate! Compared to the baseline, divorce becomes a less financially risky outcome for community property wives who increase

Figure 4.7 Counterfactual Scenario: Alternative Income Risk



(a) Risky Assets Participation Rate (b) Risky Asset Share

Note: This figure plots the property regime gap in the participation and portfolio share in risky financial assets in the counterfactual scenario and the baseline economy. The gap is computed as the difference between separate-property and community-property households' outcomes.

their demand for risky assets. This reduces the average differences in risky asset holdings between both types of couples.

Figures A4.4-A4.6 in the Appendix present the results for the gap in total savings-to-income ratio for each of the counterfactual scenarios. As can be inspected, the total savings-to-income ratio gap increases with the dissolution costs of marriage and income risk and decreases with larger income differentials in permanent income in favor of men.

4.8.2 The Role of the Dissolution Cost and Division Rules

In the model, marital property regimes introduce differences in (i) the allocation rule of marital savings between spouses and (ii) the dissolution costs of marriage. More precisely, separate-property spouses retain ownership of their individual portfolio in the event of divorce, while community-property spouses pool their savings together and each of them retains 50% of the total household portfolio. In addition, we assume that community-property couples pay a dissolution cost of marriage while separate-property couples face no cost.

We conduct two counterfactual exercises to isolate the contribution of each of these two factors (i.e. asset allocation vs. dissolution costs) on the estimated marital property regime gap. In the first scenario, we simulate the model assuming that both types of couples face the same dissolution cost of marriage (i.e.$\kappa = 10\%$). In the second scenario, we assume those married under separate property pool the assets upon divorce and divide them in half without paying any dissolution cost. Table 4.5

presents the difference in the risky investment participation rate, the share of risky assets, and the savings-to-income ratio in these two counterfactual economies with respect to the baseline for separate-property couples.

Table 4.5 Disentangling the Role of Dissolution Costs vs. Asset Allocation Rule

	(1) The role of dissolution cost	(2) The role of pooling assets
Risky assets participation rates	-7.26 p.p	-8.5 p.p
Risky assets share	-3.28 p.p	3.05 p.p
Total savings-to-income ratio	0.43 p.p	4.55 p.p

Note: Columns (1) and (2) present the percentage points difference between the model outcomes in each of the two counterfactual scenarios and the baseline for separate-property couples. In the first column, we assume that separate property couples also pay the dissolution cost, κ. In the second column, we assume that separate property couples also pool the assets and divide them by half in case of divorce.

Column (1) in Table 4.5 shows that when separate-property wives bear the same dissolution cost as community-property wives, they save more but demand less risky assets. Higher dissolution costs make divorce riskier, as a fraction of permanent income is destroyed in the event of divorce, which encourages higher precautionary savings in the form of safe assets both at the extensive and intensive margin to smooth consumption. Column (2) in Table 4.5 shows that when separate-property couples pool the assets and divide them fifty-fifty in the event of divorce, they would also save more and demand less risky assets at the extensive margin. However, they would increase their risky investment at the intensive margin (i.e. they would allocate a higher share of their portfolio to risky assets). Quantitatively, the dissolution costs of marriage seem to be more important for explaining the property regime gap in risky investment at the extensive and fully explains it at the intensive margin. Instead, the fact that assets are split equally between spouses regardless of the intra-household distribution of savings during marriage seems to be quantitatively more relevant for explaining the difference between couples in savings accumulation.

4.9 Model validation

The empirical findings presented in Section 4.4 suggest that separate-property couples hold significantly riskier portfolios than community-property ones *only* when wives take a more prominent role in managing household finances. We validate our theoretical results by solving the model when the husband is the one making portfolio choices taking as given her wife's saving decisions.

Table 4.6 presents the relevant parameters modified for this exercise and their corresponding values. Relative to the baseline economy, we change both spouses' income parameters to match the income profiles of male-headed households in the EFF data 2002-2020. In particular, we change the permanent income components to match the average gender wage gap for male-headed households and estimate the variance and persistence of the stochastic component of both spouses' income for these couples. Finally, we also obtain the wife's total savings and share in risky assets following the procedure explained in Section 4.6.[12] It is noteworthy that relative to the baseline economy, husbands leading household finances have a higher level of permanent income but a lower variance of the income shock compared to wives leading household finances. Conversely, the spouse in this case - the wife - maintains lower savings levels and a relatively smaller portfolio of risky assets.

Table 4.6 Parameters When the Husband is the Household Head

Parameter	**Value**	**Source**
$\bar{y}^h$	28305.80	EFF
$\bar{y}^w$	14819.79	EFF
σ_h^2	0.349	EFF
σ_w^2	0.297	EFF
ρ_h	0.514	EFF
ρ_w	0.574	EFF
α_1	19.15%	EFF
α_2	6.42%	EFF

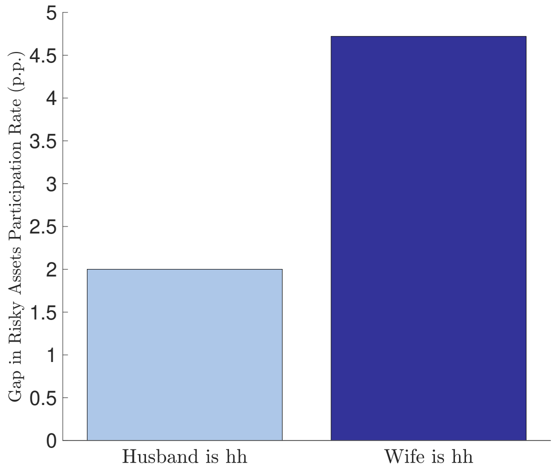
Figure 4.8 compares the gap in risky asset participation rates in this alternative economy with the baseline one. As can be inspected, the gap in risky asset participation rates shrinks by more than 2 pp when we match key moments of male-headed couples' income profiles. These results highlight the importance of income profile heterogeneity in explaining differences in portfolio investments for couples with the same property division rules.

4.10 Conclusion

A vast literature in household finance emphasizes that women are less likely to take financial risks than men because of their psychological traits (less confidence and optimism, more risk aversion) or because of the social norms they have been raised in (financial matters are considered the domain of men). This paper uncovers a critical yet unexplored determinant of financial investment when women are in charge of household finances: the marital property regime.

[12] Appendix Table A4.8 presents the parameter estimates of the income process of male-headed households, whereas Table A4.7 displays the values utilized for the wife's total savings and share in risky assets.

Figure 4.8 Model Validation



Note: This figure plots the property regime gap in the risky assets participation rate when calibrating the model to match male-headed households' income profiles and compares it with the baseline economy (female-headed households). The gap is computed as the difference between separate-property and community-property households' outcomes.

We use rich household-level data and exploit the regional variation in default marital property regimes in Spain to provide causal evidence on the effects of property division rules on couples' risky financial investment. We find that couples married under separate property are more likely to hold wealth in risky assets than their counterparts married under community property when women are in charge of household finances. Not only do these couples participate more in risky assets, but also they hold a more diversified portfolio towards risky assets. In particular, separate-property households are up to 9% more likely to take financial risks than those married under community property. On average, they also hold a share in risky asset classes up to 5 percentage points higher.

To understand better the mechanisms at play, we develop a two-period financial portfolio choice model where wives decide how to allocate savings and couples differ in their property division rule. Couples consist of two individuals born married and face an exogenous probability of divorce in the second period. In the model, property division rules determine the sharing rule of marital savings upon divorce and the associated dissolution costs of marital assets. In the event of divorce, separate-property spouses take their individual assets and face no dissolution of marital assets while community-property couples must pay the costs of dissolving the common pool of assets equally between spouses. We calibrate the model to match key moments of Spanish female-headed couples and show that divorce risk and gender differences in labor income profiles are key determinants in shaping the financial portfolio choices of married couples under different property division rules.

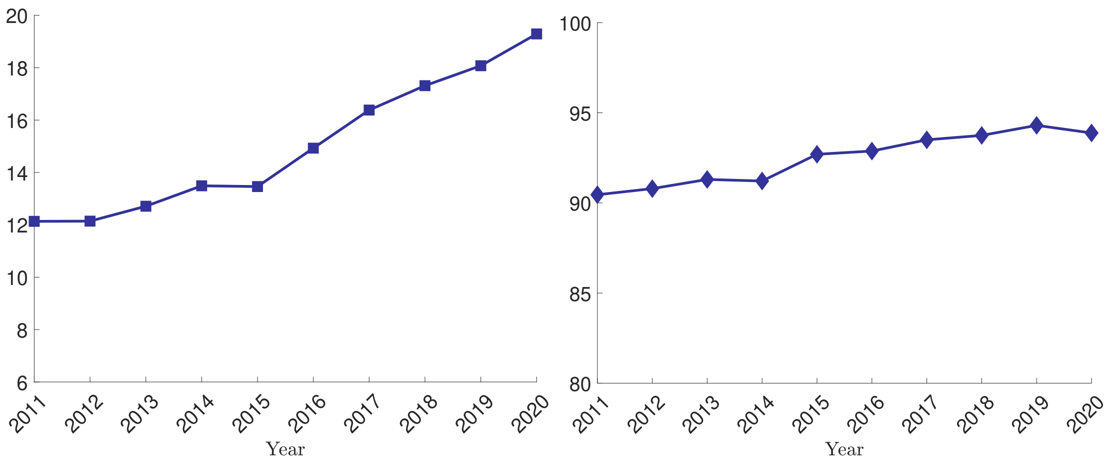
In all, our results suggest that property division rules in marriage seem to be an essential factor influencing the portfolio choices of couples in the face of divorce risk.

An exciting extension of this work would be to analyze the wealth accumulation outcomes of divorced women under these two regimes and their implications for explaining the gender wealth gap later in life. We leave this for future research.

A4 Appendix to Chapter 4

A4.1 Institutional Background

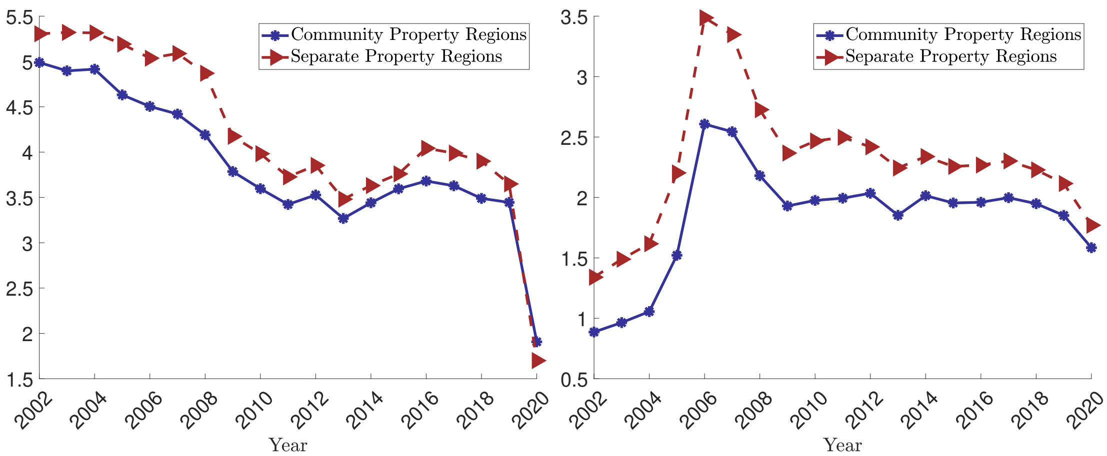
Figure A4.1 Prenuptial Contracts



(a) Prenuptial Contracts (% Marriages) (b) Separate Property (% Total Contracts)

Note: The figure plots the evolution of prenuptial contracts (% total marriages) and prenuptial contracts for separate property (% total prenuptial contracts) between 2011-2020. The data has been obtained from Statistics of the General Council of Notaries.

Figure A4.2 Marriages and Divorces in Spanish Regions by Default Regime



(a) Marriages per 1000 inhabitants (b) Divorces per 1000 inhabitants

Note: The figure plots the evolution of marriages and divorces per 1000 inhabitants across Spanish regions depending on their default property regime for the period 2002-2020. Separate-property regions (blue triangle line) are Catalonia and the Balearic Islands (and Valencian Community for the period 2009-2015). Community-property regions (red star line) are the rest of the Spanish regions (and Valencian Community for the period 2002-2008, 2016-2017).

A4.2 Household Data

Table A4.1 Household Summary Statistics - Wife is the Household Head

	Mean	St. dev.	Separate	Community
Panel A. Socioeconomic characteristics				
Household head				
Separate property	0.25	0.43		
Age	44	7.98	44	44
Education				
Less than high school	0.24	0.43	0.19	0.26
High School	0.35	0.48	0.31	0.36
College	0.40	0.49	0.50	0.37
Occupation in financial sector	0.05	0.23	0.08	0.05
Comparative ratios bw spouses				
Education ratio bw spouses	1.24	0.56	1.20	1.26
Age ratio bw spouses	0.98	0.09	0.98	0.97
Wage ratio bw spouses	0.83	0.65	0.89	0.81
Other controls				
Home-ownership				
Rent	0.11	0.29	0.13	0.11
Ownership	0.84	0.33	0.82	0.85
Other	0.04	0.18	0.05	0.04
Household size	3.55	0.99	3.52	3.56
Income (thousands eur)	55.12	46.98	67.52	51.08
Net wealth (thousands eur)	306.46	614.22	464.76	254.90
Panel B. Financial Variables				
Financial Variables				
Participation risky assets	0.22	0.41	0.33	0.18
Risky asset classes (%Total asset classes)	0.11	0.21	0.17	0.09
Risky assets share	0.10	0.24	0.16	0.08

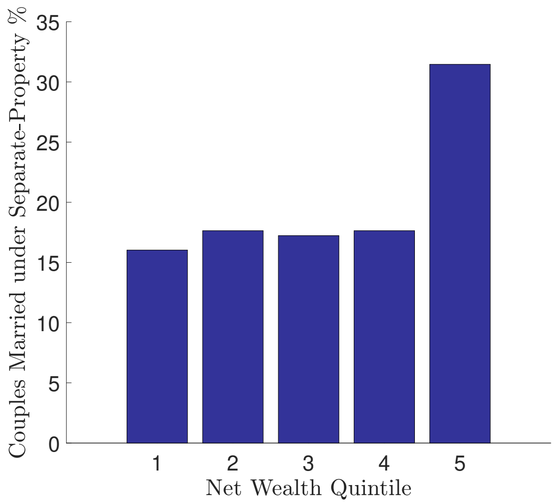
Note: This table shows summary statistics for two-spouse households characteristics and by marital property regime of the household head. The sample includes information from 2002-2020 waves of the Spanish Survey of Household Finances and is restricted to two-spouse households aged above 25 years old who are employed. Self-employed households are excluded from the sample. Observations: 1681 (1652 for the education ratio, 1633 for the risky asset classes share, and 1626 for the risky assets share).

Table A4.2 Household Summary Statistics - Husband is the Household Head

	Mean	St. dev.	Separate	Community
Panel A. Socioeconomic characteristics				
Household head				
Separate property	0.27	0.44		
Age	47	8.88	47	47
Education				
Less than high school	0.23	0.42	0.15	0.26
High School	0.33	0.47	0.31	0.34
College	0.34	0.50	0.54	0.40
Occupation in financial sector	0.05	0.22	0.09	0.04
Comparative ratios bw spouses				
Education ratio bw spouses	1.04	0.41	1.05	1.03
Age ratio bw spouses	1.06	0.09	1.06	1.06
Wage ratio bw spouses	1.98	2.09	2.14	1.92
Other controls				
Home-ownership				
Rent	0.08	0.27	0.08	0.08
Ownership	0.89	0.32	0.87	0.89
Other	0.03	0.17	0.04	0.03
Household size	3.50	1.00	3.45	3.51
Income (thousands eur)	73.17	109.00	101.28	62.92
Net wealth (thousands eur)	679.90	4186.65	1438.94	403.04
Panel B. Financial Variables				
Financial Variables				
Participation risky assets	0.35	0.48	0.41	0.32
Risky asset classes (%Total asset classes)	0.18	0.25	0.21	0.16
Risky assets share	0.18	0.31	0.23	0.16

Note: This table shows summary statistics for two-spouse households characteristics and by marital property regime of the household head. The sample includes information from the 2002-2020 waves of the Spanish Survey of Household Finances and is restricted to two-spouse households aged above 25 years old who are employed. Self-employed households are excluded from the sample. Observations: 3229 (3148 for the education ratio, 3158 for the risky asset classes share, and 3148 for the risky assets share).

Figure A4.3 Separate Property Couples in Community Property Regions



Note: The figure shows the proportion of married couples that opt out of community property by net wealth percentile as a share of total married couples opting out. Data are from the 2002-2020 waves of the Spanish Survey of Household Finances. The sample is restricted to two-earner households aged above 25. Self-employed households are excluded.

A4.3 Empirical Results

Table A4.3 Robustness Checks - Participation in Risky Financial Assets

	(1) Risky Financial Assets	(2) Risky Financial Assets	(3) Risky Financial Assets
Separate Property	-0.060	-0.086*	-0.061
	(0.036)	(0.044)	(0.040)
Female	-0.056***	-0.096***	-0.074***
	(0.016)	(0.015)	(0.015)
Female × Sep. Property	0.084***	0.151***	0.095***
	(0.022)	(0.022)	(0.023)
Risk Attitudes	✓		
Online Banking		✓	
Managed Fin. Accounts		✓	
Mother Housewife			✓
Households Characteristics	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Observations	4262	3087	4216

Note: The sample includes all two-earner married households in 2002-2020. This table reports 2SLS estimates from a model where the dependent variable is a binary variable that equals 1 if households hold wealth in risky assets. Separate property is instrumented using a dummy for residence in Catalonia or the Balearic Islands. *Female* is a dummy variable that equals 1 if the headship of the household is female and 0 otherwise. *Risk attitudes* is a categorical variable that measures attitudes towards risk from a lower to a higher degree of risk tolerance. *Online banking* is a dummy variable for online banking usage. *Managed Fin Accounts* is a dummy variable for ownership of managed financial accounts by professional financial institutions. *Mother Housewife* is a dummy variable that equals 1 if the mother of the household head is/was a housewife. We exclude from the sample couples living in Valencian Community as this region changed its default regime during the time period considered. Standard errors (in parenthesis) are robust and clustered at the regional level.

Table A4.4 Robustness Checks - Portfolio Share in Risky Asset Classes

	(1) Risky Financial Assets	(2) Risky Financial Assets	(3) Risky Financial Assets
Separate Property	-0.038*	-0.044*	-0.038
	(0.023)	(0.026)	(0.025)
Female	-0.028***	-0.049***	-0.038***
	(0.007)	(0.008)	(0.010)
Female × Sep. Property	0.047***	0.078***	0.052***
	(0.012)	(0.016)	(0.016)
Risk Attitudes	✓		
Online Banking		✓	
Managed Fin. Accounts		✓	
Mother Housewife			✓
Households Characteristics	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Observations	4156	3012	4113

Note: The sample includes all two-earner married households in 2002-2020. This table reports 2SLS estimates from a model where the dependent variable is a binary variable that equals 1 if households hold wealth in risky assets - mutual funds, listed shares, and unlisted shares. Separate property is instrumented using a dummy for residence in Catalonia or the Balearic Islands. *Female* is a dummy variable that equals 1 if the headship of the household is female and 0 otherwise. *Risk attitudes* is a categorical variable that measures attitudes towards risk from a lower to a higher degree of risk tolerance. *Online banking* is a dummy variable for online banking usage. *Managed Fin Accounts* is a dummy variable for ownership of managed financial accounts by professional financial institutions. *Mother Housewife* is a dummy variable that equals 1 if the mother of the household head is/was a housewife. We exclude from the sample couples living in Valencian Community as this region changed its default regime during the time period considered. Standard errors (in parenthesis) are robust and clustered at the regional level.

Table A4.5 Empirical Gaps

	(1) % Risky Financial Assets Assets	(2) Risky Financial Assets	(3) Savings-to-Income Ratio
	Wife household head	Wife household head	Wife household head
Separate Property	0.023	0.047**	-0.043*
	(0.017)	(0.022)	(0.020)
Households Characteristics	Yes	Yes	Yes
Survey Year FE	Yes	Yes	Yes
Observations	1461	1461	1461

Note: The sample includes all two-earner married households in 2002-2020 where the household head is the wife. This table reports 2SLS estimates from a model where the dependent variable is the share in risky financial assets in the household portfolio (column (1)), a binary variable that equals 1 if households hold wealth in risky assets (column (2)) and the ratio between savings and total household income (column (3)). Separate property is instrumented using a dummy for residence in Catalonia or the Balearic Islands. We exclude from the sample couples living in Valencian Community as this region changed its default regime during the time period considered. Standard errors (in parenthesis) are robust and clustered at the regional level.

A4.4 Model Calibration and Theoretical Results

Table A4.6 Husband Savings Calibration

Parameter		Data Source
	Married Couples Wife is the household head	
Household savings-to-income ratio	0.434	EFF
Husband savings-to-income ratio	0.241	

Note: The average household savings-to-income ratio has been computed using the panel structure of the EFF survey data from 2002-2020. We measure savings as a flow, that is, savings refers to the change in total household financial savings in two consecutive waves. Income only includes labor income. The sample has been restricted to two-earner married couples above 25 years old, for which the wive is the most knowledgeable about household finances. The gender wage gap is 1.25 for these couples. Survey weights are applied to give consistent averages for the Spanish population.

Table A4.7 Wife Savings Calibration

Parameter		Data Source
	Married Couples Husband is the household head	
Household savings-to-income ratio	0.557	EFF
Wife savings-to-income ratio	0.192	
Wife share in risky assets	0.064	EFF

Note: The average household savings-to-income ratio has been computed using the panel structure of the EFF survey data from 2002-2020. We measure savings as a flow, that is, savings refers to the change in total household financial savings in two consecutive waves. Income only includes labor income. The sample has been restricted to two-earner married couples above 25 years old, for which the husband is the most knowledgeable about household finances. The gender wage gap is 1.91 for these couples. The risky share for wives has been computed using a sample of divorced women in the same period. Survey weights are applied to give consistent averages for the Spanish population.

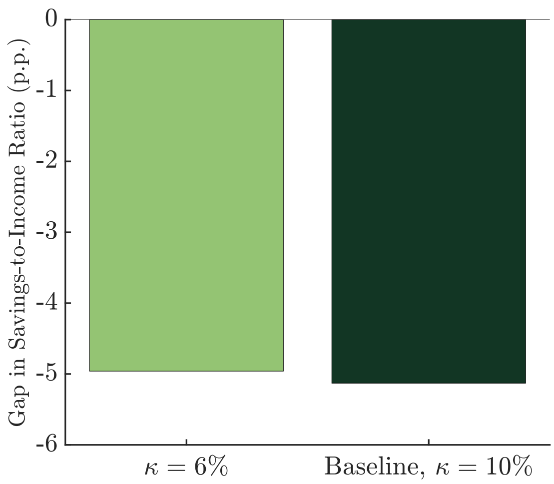
Table A4.8 Estimation Results - Stochastic Income Process

Parameter	Married Couples
	Husband is the household head
σ_h^2	0.349
ρ_h	0.514
σ_w^2	0.297
ρ_w^2	0.574
$\bar{y}^h$	28305.80
$\bar{y}^w$	14819.79

Table A4.9 Counterfactual - Divorce risk

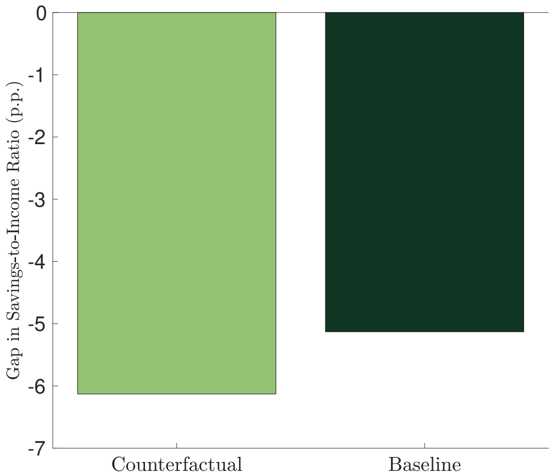
	(1) Baseline	(2) Counterfactual	(3) Data
Gap in	$\delta = 0.24$	$\delta = 0$	
Risky assets share	1.5 p.p	0 p.p	2.3 p.p
Risky assets participation rates	4.7 p.p	0 p.p	4.7 p.p
Total savings-to-income ratio	-5.1 p.p	0 p.p	-4.3 p.p

Figure A4.4 Gap in Savings-to-income Ratio - Alternative Dissolution Costs of Marriage



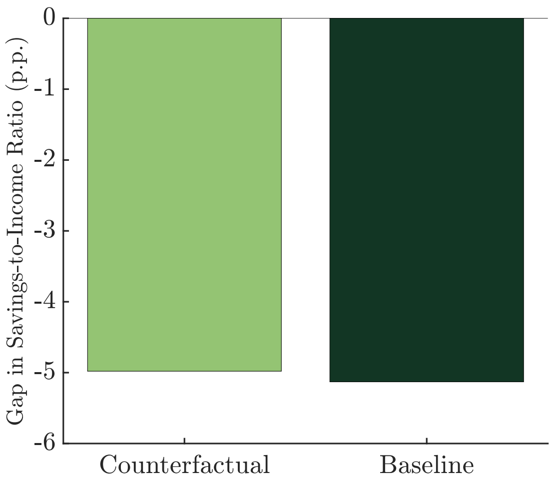
Note: This figure plots the property regime gap in the total savings-to-income ratio generated by the model in the baseline economy and counterfactual scenario. The gap is computed as the difference between separate-property and community-property households' outcomes.

Figure A4.5 Gap in Savings-to-income Ratio - Alternative Income Levels



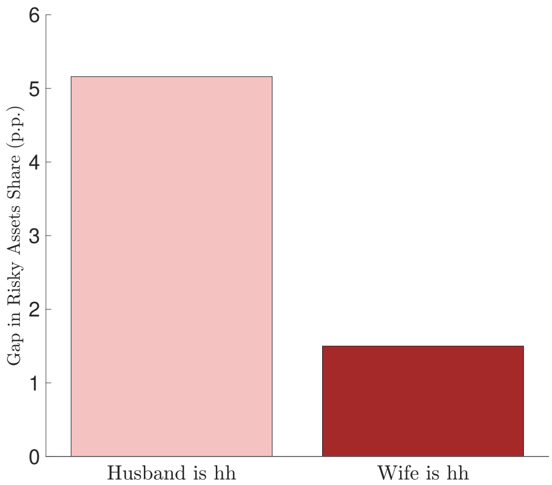
Note: This figure plots the property regime gap in the total savings-to-income ratio generated by the model in the baseline economy and counterfactual scenario. The gap is computed as the difference between separate-property and community-property households' outcomes.

Figure A4.6 Gap in Savings-to-income Ratio - Alternative Income Risk



Note: This figure plots the property regime gap in the total savings-to-income ratio generated by the model in the baseline economy and counterfactual scenario. The gap is computed as the difference between separate-property and community-property households' outcomes.

Figure A4.7 Total savings-to-income Ratio



(a) Risky Assets Share

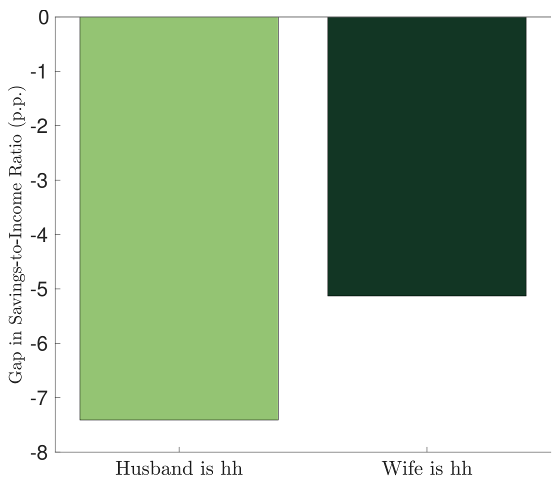


Figure A4.8 Model Validation - Property Regime Gaps for Male-headed Households

Note: This figure plots the property regime gap in the portfolio share in risky financial assets and the total savings-to-income ratio generated by the model, and the one estimated in the data. The gap is computed as the difference between separate-property and community-property households' outcomes. The darker bar refers to the 2SLS estimate of the gap and the corresponding 95% CI using EFF survey waves 2002-2020. The lighter bar refers to the model simulation outcome.

Bibliography

Abraham, K. G. and Kearney, M. S. (2020). Explaining the decline in the US employment-to-population ratio: A review of the evidence. *Journal of Economic Literature*, 58(3):585–643.

Abramowitz, J. (2017). Assisted reproductive technology and women's timing of marriage and childbearing. *Journal of Family and Economic Issues*, 38(1):100–117.

Adda, J., Dustmann, C., and Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2):293–337.

Adserà, A. (2004). Changing fertility rates in developed countries. the impact of labor market institutions. *Journal of Population Economics*, 17(1):17–43.

Ahn, N. and Sánchez-Marcos, V. (2020). Analysis of fertility using cohort-specific socio-economic data. *Review of Economics of the Household*, 18(3):711–733.

Aksoy, Y., Basso, H. S., Smith, R. P., and Grasl, T. (2019). Demographic structure and macroeconomic trends. *American Economic Journal: Macroeconomics*, 11(1):193–222.

Albanesi, S. and Olivetti, C. (2016). Gender roles and medical progress. *Journal of Political Economy*, 124(3):650–695.

Albanesi, S. and Prados, M. J. (2022). Slowing women's labor force participation: The role of income inequality. Technical report, National Bureau of Economic Research.

Alesina, A., Ichino, A., and Karabarbounis, L. (2011). Gender-based taxation and the division of family chores. *American Economic Journal: Economic Policy*, 3(2):1–40.

Amuedo-Dorantes, C. and Kimmel, J. (2005). The motherhood wage gap for women in the United States: The importance of college and fertility delay. *Review of Economics of the Household*, 3(1):17–48.

Anghel, B., Conde-Ruiz, J. I., and De Artíñano, I. M. (2019). Brechas salariales de género en España. *Hacienda Pública Española*, (229):87–119.

Arroyo, C. R. and Zhang, J. (1997). Dynamic microeconomic models of fertility choice: A survey. *Journal of Population Economics*, 10(1):23–65.

Asai, Y. (2015). Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from Japan. *Labour Economics*, 36:72–83.

Attanasio, O., Kitao, S., and Violante, G. L. (2007). Global demographic trends and social security reform. *Journal of Monetary Economics*, 54(1):144–198.

Attanasio, O., Low, H., and Sánchez-Marcos, V. (2008). Explaining changes in female labor supply in a life-cycle model. *American Economic Review*, 98(4):1517–52.

Auer, W. and Danzer, N. (2016). Fixed-term employment and fertility: Evidence from German micro data. *CESifo Economic Studies*, 62(4):595–623.

Auerbach, A. J. and Kotlikoff, L. J. (1987). Evaluating fiscal policy with a dynamic simulation model. *The American Economic Review*, 77(2):49–55.

Bacher, A. (2021a). The gender investment gap over the life-cycle.

Bacher, A. (2021b). Housing and savings behavior across family types.

Bajtelsmit, V. L. and Bernasek, A. (1996). Why do women invest differently than men? *Financial Counseling and Planning*, 7.

Bankole, A. (1995). Desired fertility and fertility behaviour among the yoruba of nigeria: A study of couple preferences and subsequent fertility. *Population Studies*, 49(2):317–328.

Barber, B. M. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1):261–292.

Becker, G. S. (1960). An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries*, NBER Chapters, pages 209–240. National Bureau of Economic Research, Inc.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2):S279–S288.

Bentolila, S., Cahuc, P., Dolado, J. J., and Barbanchon, T. L. (2012). Two-tier labour markets in the great recession: France versus Spain. *The Economic Journal*, 122(562):F155–F187.

Bertocchi, G., Brunetti, M., and Torricelli, C. (2011). Marriage and other risky assets: A portfolio approach. *Journal of Banking & Finance*, 35(11):2902–2915.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–55.

Bettendorf, L. J., Jongen, E. L., and Muller, P. (2015). Childcare subsidies and labour supply — evidence from a large Dutch reform. *Labour Economics*, 36:112–123.

Bick, A. (2016). The quantitative role of child care for female labor force participation and fertility. *Journal of the European Economic Association*, 14(3):639–668.

Bick, A., Brüggemann, B., Fuchs-Schündeln, N., and Paule-Paludkiewicz, H. (2019). Long-term changes in married couples' labor supply and taxes: Evidence from the US and Europe since the 1980s. *Journal of International Economics*, 118:44–62.

Bick, A. and Fuchs-Schündeln, N. (2017). Quantifying the disincentive effects of joint taxation on married women's labor supply. *American Economic Review*, 107(5):100–104.

Bitler, M. P. and Schmidt, L. (2012). Utilization of infertility treatments: the effects of insurance mandates. *Demography*, 49(1):125–149.

Blackburn, M. L., Bloom, D. E., and Neumark, D. (1993). Fertility timing, wages, and human capital. *Journal of Population Economics*, 6(1):1–30.

Blanco, M. A., Bauluz, L., and Martínez-Toledano, C. (2021). Wealth in Spain 1900–2017 a country of two lands. *The Economic Journal*, 131(633):129–155.

Blankenau, W., Kose, M. A., and Yi, K.-M. (2001). Can world real interest rates explain business cycles in a small open economy? *Journal of Economic Dynamics and Control*, 25(6-7):867–889.

Blau, D. M. and Robins, P. K. (1988). Child-care costs and family labor supply. *The Review of Economics and Statistics*, 70(3):374–381.

Blau, F. D. and Kahn, L. M. (2013). Female labor supply: Why is the United States falling behind? *American Economic Review*, 103(3):251–56.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Bongardt, A. and Torres, F. (2020). The Lisbon strategy and Europe 2020. In *Oxford Research Encyclopedia of Politics*.

Bonin, H., Gil, J., and Patxot, C. (2001). Beyond the Toledo agreement: the intergenerational impact of the Spanish pension reform. *Spanish Economic Review*, 3:111–130.

Borella, M., De Nardi, M., Pak, M., Russo, N., and Yang, F. (2022). U.S. taxes effective over time. Technical report, Working Paper.

Borella, M., De Nardi, M., and Yang, F. (2019). Are marriage-related taxes and social security benefits holding back female labor supply? Technical report, National Bureau of Economic Research.

Boskin, M. J. and Sheshinski, E. (1983). Optimal tax treatment of the family: Married couples. *Journal of Public Economics*, 20(3):281–297.

Brassiolo, P. (2013). The effect of property division laws on divorce and labor supply: evidence from Spain.

Bronson, M. A. and Mazzocco, M. (2018). Taxation and household decisions: an intertemporal analysis. Technical report, Working Paper.

Bucher-Koenen, T., Lusardi, A., Alessie, R., and Van Rooij, M. (2017). How financially literate are women? an overview and new insights. *Journal of Consumer Affairs*, 51(2):255–283.

Buckles, K. (2007). Stopping the biological clock: Infertility treatments and the career-family tradeoff. *Boston University, Unpublished Manuscript*.

Buttet, S. and Schoonbroodt, A. (2013). An accounting exercise for the shift in life-cycle employment profiles of married women born between 1940 and 1960. *Journal for Labour Market Research*, 46(3):253–271.

Butz, W. P. and Ward, M. P. (1979). The emergence of countercyclical U.S. fertility. *American Economic Review*, 69(3):318–328.

Bünning, M. (2015). What happens after the 'daddy months'? fathers' involvement in paid work, childcare, and housework after taking parental leave in German. *European Sociological Review*, 31(6):738–748.

Cascio, E. U., Haider, S. J., and Nielsen, H. S. (2015). The effectiveness of policies that promote labor force participation of women with children: A collection of national studies. *Labour Economics*, 36:64 – 71.

Caucutt, E. M., Guner, N., and Knowles, J. (2002). Why do women wait? matching, wage inequality, and the incentives for fertility delay. *Review of Economic Dynamics*, 5(4):815–855.

Chiappori, P.-A., Fortin, B., and Lacroix, G. (2002). Marriage market divorce legislation and household labor supply. *Journal of Political Economy*, 110(1):37–72.

Cigno, A. and Ermisch, J. (1989). A microeconomic analysis of the timing of births. *European Economic Review*, 33(4):737–760.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, 18(2):491–533.

Cohen, A., Dehejia, R., and Romanov, D. (2013). Financial incentives and fertility. *The Review of Economics and Statistics*, 95(1):1–20.

Conde-Ruiz, J. I. and González, C. I. (2015). Challenges for Spanish pensions in the early 21st century. *CESifo DICE Report*, 13(2):20–24.

Conesa, J. C., Kehoe, T. J., Nygaard, V. M., and Raveendranathan, G. (2020). Implications of increasing college attainment for aging in general equilibrium. *European Economic Review*, 122:103363.

Conesa, J. C. and Krueger, D. (2006). On the optimal progressivity of the income tax code. *Journal of Monetary Economics*, 53(7):1425–1450.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–74.

Cubeddu, L. and Ríos-Rull, J.-V. (2003). Families as shocks. *Journal of the European Economic Association*, 1(2-3):671–682.

Da Rocha, J. M. and Fuster, L. (2006). Why are fertility rates and female employment ratios positively correlated across oecd countries? *International Economic Review*, 47(4):1187–1222.

De la Escosura, L. P. (2017). Spain's performance in comparative perspective. In *Spanish Economic Growth, 1850–2015*, pages 39–46. Springer.

De la Fuente, Á. (2017). Series enlazadas de algunos agregados económicos regionales, 1955-2014. *FEDEA, Estudios sobre la Economía Española*, 21.

De La Fuente, Á. and Doménech, R. (2013). The financial impact of Spanish pension reform: A quick estimate. *Journal of Pension Economics & Finance*, 12(1):111–137.

De la Rica, S. and Iza, A. (2005). Career planning in Spain: Do fixed-term contracts delay marriage and parenthood? *Review of Economics of the Household*, 3(1):49–73.

De La Roca, J. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142.

De Nardi, M., French, E., Jones, J. B., and McGee, R. (2021). Why do couples and singles save during retirement? Technical report, National Bureau of Economic Research.

De Nardi, M., Imrohoroğlu, S., and Sargent, T. J. (1999). Projected US demographics and social security. *Review of Economic dynamics*, 2(3):575–615.

De Quinto, A., Hospido, L., and Sanz, C. (2021). The child penalty: evidence from Spain. *SERIEs: Journal of the Spanish Economic Association*, 12(4):585–606.

Del Boca, D. and Sauer, R. M. (2009). Life cycle employment and fertility across institutional environments. *European Economic Review*, 53(3):274–292.

Díaz-Giménez, J. and Díaz-Saavedra, J. (2009). Delaying retirement in Spain. *Review of Economic dynamics*, 12(1):147–167.

Díaz-Gimenez, J. and Díaz-Saavedra, J. (2017). The future of Spanish pensions. *Journal of Pension Economics & Finance*, 16(2):233–265.

Díaz Giménez, J. and Pijoan-Mas, J. (2006). Flat tax reforms in the US: a boon for the income poor.

Díaz-Saavedra, J. (2022). Heterogeneity in longevity, redistribution, and pension reform. *Journal of Pension Economics & Finance*, pages 1–36.

Díaz-Saavedra, J., Marimon, R., and de Sousa, J. B. (2023). A worker's backpack as an alternative to payg pension systems. *Journal of the European Economic Association*, page jvad021.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2022a). The economics of fertility: A new era. Technical report, National Bureau of Economic Research.

Doepke, M., Hannusch, A., Kindermann, F., and Tertilt, M. (2022b). The economics of fertility: A new era. Working Paper 29948, National Bureau of Economic Research.

Doepke, M. and Kindermann, F. (2019). Bargaining over babies: Theory, evidence, and policy implications. *American Economic Review*, 109(9):3264–3306.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dolado, J. J., Lalé, E., and Siassi, N. (2021). From dual to unified employment protection: Transition and steady state. *Quantitative Economics*, 12(2):547–585.

Duran-Franch, J. (2021). Oh, man! what happened to women? the blurring of gender-based occupational segregation.

Eckstein, Z., Keane, M., and Lifshitz, O. (2019). Career and family decisions: Cohorts born 1935–1975. *Econometrica*, 87(1):217–253.

Eckstein, Z. and Wolpin, K. I. (1989). Dynamic labour force participation of married women and endogenous work experience. *The Review of Economic Studies*, 56(3):375–390.

Ekberg, J., Eriksson, R., and Friebel, G. (2013). Parental leave — a policy evaluation of the swedish "daddy-month" reform. *Journal of Public Economics*, 97:131–143.

Erosa, A., Fuster, L., and Kambourov, G. (2016a). Towards a micro-founded theory of aggregate labour supply. *The Review of Economic Studies*, 83(3):1001–1039.

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2017). Hours, occupations, and gender differences in labor market outcomes. *NBER Working Paper Series.*

Erosa, A., Fuster, L., Kambourov, G., and Rogerson, R. (2022). Hours, occupations, and gender differences in labor market outcomes. *American Economic Journal: Macroeconomics*, 14(3):543–90.

Erosa, A., Fuster, L., and Restuccia, D. (2002). Fertility decisions and gender differences in labor turnover, employment, and wages. *Review of Economic Dynamics*, 5(4):856–891.

Erosa, A., Fuster, L., and Restuccia, D. (2010). A General Equilibrium Analysis of Parental Leave Policies. *Review of Economic Dynamics*, 13(4):742–758.

Erosa, A., Fuster, L., and Restuccia, D. (2016b). A quantitative theory of the gender gap in wages. *European Economic Review*, 85:165–187.

Esteve, A. and Cortina, C. (2006). Changes in educational assortative mating in contemporary Spain. *Demographic Research*, 14:405–428.

Farré, L. and González, L. (2019). Does paternity leave reduce fertility? *Journal of Public Economics*, 172:52–66.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500.

Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77.

Fogli, A. and Veldkamp, L. (2011). Nature or nurture? learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138.

Fortin, N. M. (2015). Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (117/118):379–401.

Francesconi, M. (2002). A joint dynamic model of fertility and work of married women. *Journal of labor Economics*, 20(2):336–380.

Frémeaux, N. and Leturcq, M. (2020). Inequalities and the individualization of wealth. *Journal of Public Economics*, 184:104145.

Fuster, L., İmrohoroğlu, A., and İmrohoroğlu, S. (2007). Elimination of social security in a dynastic framework. *The Review of Economic Studies*, 74(1):113–145.

Galasso, V. and Profeta, P. (2004). Lessons for an ageing society: the political sustainability of social security systems. *Economic Policy*, 19(38):64–115.

Garcia-Mandicó, S. and Jiménez-Martín, S. (2020). Spillovers in pension incentives and the joint retirement behavior of Spanish couples. *Documento de Trabajo*, page 13.

García-Miralles, E., Guner, N., and Ramos, R. (2019). The Spanish personal income tax: facts and parametric estimates. *SERIEs: Journal of the Spanish Economic Association*, 10(3):439–477.

Gayle, G.-L. and Golan, L. (2012). Estimating a dynamic adverse-selection model: Labour-force experience and the changing gender earnings gap 1968–1997. *The Review of Economic Studies*, 79(1):227–267.

Geyer, J., Haan, P., and Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics*, 36:84–98.

Gipson, J. D. and Hindin, M. J. (2009). The effect of husbands' and wives' fertility preferences on the likelihood of a subsequent pregnancy, bangladesh 1998–2003. *Population Studies*, 63(2):135–146. PMID: 19536696.

Givord, P. and Marbot, C. (2015). Does the cost of child care affect female labor market participation? an evaluation of a french reform of childcare subsidies. *Labour Economics*, 36:99–111.

Global Gender Gap Report, W. E. F. (2022). Global gender gap report 2022.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2):1–21.

Goldin, C. (2021). *Career and Family: Women's Century-Long Journey toward Equity*. Princeton University Press.

Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, 110(4):730–770.

Goldin, C. and Katz, L. F. (2008). Transitions: Career and family life cycles of the educational elite. *American Economic Review*, 98(2):363–69.

Goldin, C. and Mitchell, J. (2017). The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives*, 31(1):161–82.

Gomes, F., Haliassos, M., and Ramadorai, T. (2021). Household finance. *Journal of Economic Literature*, 59(3):919–1000.

González, L. and Trommlerová, S. K. (2023). Cash transfers and fertility: How the introduction and cancellation of a child benefit affected births and abortions. *Journal of Human Resources*, 58(3):783–818.

González, L. (2013). The effect of a universal child benefit on conceptions, abortions, and early maternal labor supply. *American Economic Journal: Economic Policy*, 5(3):160–188.

Grabka, M. M., Marcus, J., and Sierminska, E. (2015). Wealth distribution within couples. *Review of Economics of the Household*, 13:459–486.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. *American Economic Review*, 104(5):348–53.

Greenwood, J., Seshadri, A., and Yorukoglu, M. (2005). Engines of liberation. *The Review of Economic Studies*, 72(1):109–133.

Groes, F., Iorio, D., Leung, M. Y., Santaeulalia-Llopis, R., et al. (2017). Educational disparities in the battle against infertility: Evidence from ivf success. *Copenhagen Business School, Unpublished Manuscript*.

Guiso, L. and Zaccaria, L. (2021). From patriarchy to partnership: Gender equality and household finance. Technical report, Einaudi Institute for Economics and Finance (EIEF).

Guner, N., Kaya, E., and Sánchez-Marcos, V. (2014). Gender gaps in Spain: policies and outcomes over the last three decades. *SERIEs*, 5(1):61–103.

Guner, N., Kaya, E., and Sánchez Marcos, V. (2021). Labor market institutions and fertility. BSE Working Paper 1297.

Guner, N., Kaygusuz, R., and Ventura, G. (2012a). Taxation and household labour supply. *The Review of Economic Studies*, 79(3):1113–1149.

Guner, N., Kaygusuz, R., and Ventura, G. (2012b). Taxing women: A macroeconomic analysis. *Journal of Monetary Economics*, 59(1):111–128.

Guner, N., Kaygusuz, R., and Ventura, G. (2020). Child-related transfers, household labour supply, and welfare. *The Review of Economic Studies*, 87(5):2290–2321.

Haeck, C., Lefebvre, P., and Merrigan, P. (2015). Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics*, 36:137–157.

Hamilton, B. H. and McManus, B. (2012). The effects of insurance mandates on choices and outcomes in infertility treatment markets. *Health Economics*, 21(8):994–1016.

Hannusch, A. et al. (2019). Taxing families: The impact of child-related transfers on maternal labor supply. *Unpublished Manuscript, Mannheim University.*

Heckman, J. and Walker, J. (1990). The relationship between wages and income and the timing and spacing of births: Evidence from Swedish longitudinal data. *Econometrica*, 58(6):1411–41.

Heckman, J. J. and Macurdy, T. E. (1980). A life cycle model of female labour supply. *The Review of Economic Studies*, 47(1):47–74.

Hener, T. (2015). Do couples bargain over fertility? *Annals of Economics and Statistics*, (117/118):211–231.

Hospido, L., Izquierdo Martínez, S., and Machelett, M. (2021). The gender gap in financial competences. *Banco de España Analytical Article*, Economic Bulletin 1/2021.

Hotz, V. J., Klerman, J. A., and Willis, R. (1993). The economics of fertility in developed countries. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*, volume 1, Part A, chapter 07, pages 275–347. Elsevier, 1 edition.

Hotz, V. J. and Miller, R. A. (1988). An empirical analysis of life cycle fertility and female labor supply. *Econometrica*, 56(1):91–118.

Huang, Y., Pantano, J., Ye, H., and Yi, J. (2021). Property division upon divorce and household decisions. *Journal of Human Resources*, pages 0519–10243R3.

Hubener, A., Maurer, R., and Mitchell, O. S. (2016). How family status and social security claiming options shape optimal life cycle portfolios. *The Review of Financial Studies*, 29(4):937–978.

Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6):1255–1294.

Hyslop, D. R. (2001). Rising U.S. earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *American Economic Review*, 91(4):755–777.

Immervoll, H. and Barber, D. (2006). Can parents afford to work? childcare costs, tax-benefit policies and work incentives.

Imre, B. (2022). Essays in health and labour economics.

İmrohoroglu, A., Imrohoroglu, S., and Joines, D. H. (1995). A life cycle analysis of social security. *Economic Theory*, 6(1):83–114.

Iskhakov, F., Jørgensen, T. H., Rust, J., and Schjerning, B. (2017). The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics*, 8(2):317–365.

Jones, C. I. (2022). The end of economic growth? unintended consequences of a declining population. *American Economic Review*, 112(11):3489–3527.

Jones, L. E., Manuelli, R. E., and McGrattan, E. R. (2015). Why are married women working so much? *Journal of Demographic Economics*, 81(1):75–114.

Juhn, C. and Potter, S. (2006). Changes in labor force participation in the United States. *Journal of Economic Perspectives*, 20(3):27–46.

Kaplan, G. and Violante, G. L. (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics*, 2(4):53–87.

Kaygusuz, R. (2010). Taxes and female labor supply. *Review of Economic Dynamics*, 13(4):725–741.

Kaygusuz, R. (2015). Social security and two-earner households. *Journal of Economic Dynamics and Control*, 59:163–178.

Ke, D. (2021). Who wears the pants? gender identity norms and intrahousehold financial decision-making. *The Journal of Finance*, 76(3):1389–1425.

Keane, M. P. and Sauer, R. M. (2009). Classification error in dynamic discrete choice models: Implications for female labor supply behavior. *Econometrica*, 77(3):975–991.

Keane, M. P. and Wolpin, K. I. (2010). The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women. *International Economic Review*, 51(3):851–892.

Kitao, S. (2015). Fiscal cost of demographic transition in Japan. *Journal of Economic Dynamics and Control*, 54:37–58.

Klapper, L. and Lusardi, A. (2020). Financial literacy and financial resilience: Evidence from around the world. *Financial Management*, 49(3):589–614.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109:122–26.

Kleven, H., Landais, C., and Søgaard, J. E. (2019b). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Kohler, H.-P., Billari, F. C., and Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, 28(4):641–680.

Kotlikoff, L. J., Smetters, K., and Walliser, J. (2001). Finding a way out of america's demographic dilemma. Technical report, National Bureau of Economic Research.

Krueger, A. B. (2017). Where have all the workers gone? an inquiry into the decline of the US labor force participation rate. *Brookings Papers on Economic Activity*, 2017(2):1.

Krueger, D. and Ludwig, A. (2007). On the consequences of demographic change for rates of returns to capital, and the distribution of wealth and welfare. *Journal of Monetary Economics*, 54(1):49–87.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849.

Lalive, R. and Zweimüller, J. (2009). How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments*. *The Quarterly Journal of Economics*, 124(3):1363–1402.

LaLumia, S. (2008). The effects of joint taxation of married couples on labor supply and non-wage income. *Journal of Public Economics*, 92(7):1698–1719.

Lebedinski, L., Pavlović, D., and Vujić, S. (2021). *Applied Economics Conference: Labour, Health, Education, and Welfare: book of abstracts*. Institute of Economic Sciences.

Licandro, O., Puch, L. A., and Tamarit, J. R. R. (1998). *Crecimiento óptimo, depreciación endógena y subutilización del capital*. Instituto Valenciano de Investigaciones Económicas.

Lopes, M. (2020). Job security and fertility decisions. Available at SSRN: `https://ssrn.com/abstract=3543204` or `http://dx.doi.org/10.2139/ssrn.3543204`.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. *The Review of Financial Studies*, 23(1):385–432.

Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of economic literature*, 52(1):5–44.

Machado, M. P. and Sanz-de Galdeano, A. (2015). Coverage of infertility treatment and fertility outcomes. *SERIEs*, 6(4):407–439.

Martín, A. R. S. (2010). Endogenous retirement and public pension system reform in Spain. *Economic Modelling*, 27(1):336–349.

Mazzocco, M. (2005). Individual rather than household euler equations: Identification and estimation of individual preferences using household data. *Available at SSRN 695285*.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*. Academic Press, New York.

McGrattan, E. R. and Rogerson, R. (2008). Changes in the distribution of family hours worked since 1950. In *Frontiers of Family Economics*, volume 1, pages 115–138. Emerald Group Publishing Limited.

Meriküll, J., Kukk, M., and Rõõm, T. (2021). What explains the gender gap in wealth? evidence from administrative data. *Review of Economics of the Household*, 19:501–547.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics*, 24(3):1071–1100.

Milligan, K. (2005). Subsidizing the stork: New evidence on tax incentives and fertility. *The Review of Economics and Statistics*, 87(3):539–555.

Moffitt, R. (1984). Profiles of fertility, labour supply and wages of married women: A complete life-cycle model. *The Review of Economic Studies*, 51(2):263–278.

Moreno-Maldonado, A. and Santamaria, C. (2022). Delayed childbearing and urban revival.

Müller, K.-U. and Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? evidence from a large-scale expansion of early childcare. *Labour Economics*, 62:101776.

Nishiyama, S. (2015). Fiscal policy effects in a heterogeneous-agent olg economy with an aging population. *Journal of Economic Dynamics and Control*, 61:114–132.

Nishiyama, S. (2019). The joint labor supply decision of married couples and the US social security pension system. *Review of Economic Dynamics*, 31:277–304.

Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain. *Labour Economics*, 36:124–136.

OECD (2017). Old-age dependency ratio. *Pensions at a Glance 2017: OECD and G20 Indicators*.

Olivetti, C. (2006). Changes in women's hours of market work: The role of returns to experience. *Review of Economic Dynamics*, 9(4):557–587.

Olivetti, C. and Petrongolo, B. (2017). The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries. *Journal of Economic Perspectives*, 31(1):205–230.

Permanyer, I., Spijker, J., Blanes, A., and Renteria, E. (2018). Longevity and lifespan variation by educational attainment in Spain: 1960–2015. *Demography*, 55(6):2045–2070.

Piazzalunga, D. (2016). Till money us do part: Property division at divorce and married couples' time use behaviour. Technical report, Working paper. 41, 45, 65.

Ramey, G. and Ramey, V. A. (2009). The rug rat race. Technical report, National Bureau of Economic Research.

Rasul, I. (2008). Household bargaining over fertility: Theory and evidence from malaysia. *Journal of Development Economics*, 86(2):215–241.

Rosen, H. S. (1977). Is it time to abandon joint filing? *National Tax Journal*, pages 423–428.

Rosenzweig, M. R. and Schultz, T. P. (1985). The demand for and supply of births: Fertility and its life cycle consequences. *The American Economic Review*, 75(5):992–1015.

Sánchez-Marcos, V. (2003). Women's employment and fertility in Spain over the last twenty years. Technical report, Centro de Altisimos Estudios Rios Perez (CAERP).

Sánchez-Marcos, V. and Bethencourt, C. (2018). The effect of public pensions on women's labor market participation over a full life cycle. *Quantitative Economics*, 9(2):707–733.

Sánchez Martín, A. R. and Sánchez-Marcos, V. (2010). Demographic change and pension reform in Spain: An assessment in a two-earner, olg model. *Fiscal studies*, 31(3):405–452.

Schmidt, L. (2007). Effects of infertility insurance mandates on fertility. *Journal of Health Economics*, 26(3):431–446.

Shang, Q. and Weinberg, B. A. (2013). Opting for families: Recent trends in the fertility of highly educated women. *Journal of Population Economics*, 26(1):5–32.

Sheran, M. (2007). The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions. *Review of Economic Dynamics*, 10(3):367–399.

Sommer, K. (2016). Fertility choice in a life cycle model with idiosyncratic uninsurable earnings risk. *Journal of Monetary Economics*, 83:27–38.

Tamm, M. (2019). Fathers' parental leave-taking, childcare involvement and labor market participation. *Labour Economics*, 59:184–197. Special Issue on "European Association of Labour Economists, 30th annual conference, Lyon, France, 13-15 September 2018.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters*, 20(2):177–181.

Tauchen, G. and Hussey, R. (1991). Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–396.

Testa, M. R., Cavalli, L., and Rosina, A. (2014). The effect of couple disagreement about child-timing intentions: A parity-specific approach. *Population and Development Review*, 40(1):31–53.

Thomson, E. (1997). Couple childbearing desires, intentions, and births. *Demography*, 34(3):343–354.

Thomson, E. and Hoem, J. M. (1998). Couple childbearing plans and births in sweden. *Demography*, 35(3):315–322.

Trussell, J. and Wilson, C. (1985). Sterility in a population with natural fertility. *Population Studies*, 39(2):269–286.

United Nations, Department of Economic and Social Affairs, Population Division (2022). World population prospects. Online Edition.

Van der Klaauw, W. (1996). Female labour supply and marital status decisions: A life-cycle model. *The Review of Economics Studies*, 63(2):199–235.

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2):449–472.

Ventura, G. (1999). Flat tax reform: A quantitative exploration. *Journal of Economic Dynamics and Control*, 23(9-10):1425–1458.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review*, 105(8):2295–2332.

Ward, M. P. and Butz, W. P. (1980). Completed Fertility and Its Timing. *Journal of Political Economy*, 88(5):917–940.

Westoff, C. F. (2010). Desired number of children: 2000-2008. Comparative Reports 25, DHS.

Willis, R. J. (1973). A new approach to the economic theory of fertility behavior. *Journal of Political Economy*, 81(2):S14–S64.

Wolpin, K. I. (1984). An estimable dynamic stochastic model of fertility and child mortality. *Journal of Political Economy*, 92(5):852–874.

Yamaguchi, S., Ruiz, C., Mazzocco, M., et al. (2014). Labor supply, wealth dynamics and marriage decisions. In *2014 Meeting Papers*, number 210. Society for Economic Dynamics.

Yang, Y. and Morgan, S. P. (2003). How big are educational and racial fertility differentials in the us? *Social Biology*, 50(3-4):167–187.

Zeman, K., Brzozowska, Z., Sobotka, T., Beaujouan, É., and Matysiak, A. (2014). Cohort fertility and education database. methods protocol. Online; accessed 23 January 2020.